National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

VALIDATION OF CLINICAL OUTCOME MEASURES FOR PATELLOFEMORAL PAIN SYNDROME

BY

ELIZABETH L. HARRISON

©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

DEPARTMENT OF PHYSICAL EDUCATION AND SPORT STUDIES

EDMONTON, ALBERTA
FALL 1994

National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-95199-0

Canada

UNIVERSITY OF ALBERTA

RELEASE FORM

NAME OF AUTHOR: **ELIZABETH L. HARRISON**

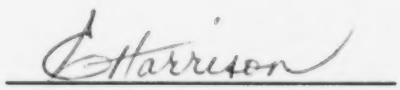
TITLE OF THESIS: **VALIDATION OF CLINICAL OUTCOME MEASURES FOR PATELLOFEMORAL PAIN SYNDROME**

DEGREE: **DOCTOR OF PHILOSOPHY**

YEAR THIS DEGREE GRANTED: **1994**

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as hereinbefore provided neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.



1520 Cairns Ave.
Saskatoon, Saskatchewan
S7H 2H6

Date: July 18/94

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Validation of Clinical Outcome Measures for Patellofemoral Pain Syndrome** submitted by **Elizabeth L. Harrison** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy**.

Dr. H. A. Quinney

Dr. D. Magee

Dr. D. Reid

Dr. D. Marshall

Dr. J. Vargo

Dr. R. Steadward

Dr. J. Taunton

Date: June 20/94

DEDICATION

I would like to dedicate this work to my parents who taught me to be myself, work hard, strive for excellence and above all to enjoy life. Their influences in my life have allowed me to pursue activities that have facilitated my growth and provided me with many challenges.

ABSTRACT

Patellofemoral Pain Syndrome (PFPS), commonly found in adolescent and young adult populations, affects individuals' lifestyles due to the symptoms of peri-patellar pain during functional activities. Limited PFPS outcome measures have resulted in an inability to appropriately monitor patients and determine the effectiveness of treatment and prevention programs. The goal of these studies was to develop a valid clinical tool that would be effective for evaluative research and clinical practice.

There were four stages in the development of the PFPS outcome measures. Stage one consisted of investigation of psychometric properties of six evaluation components: a functional index questionnaire (FIQ), visual analogue scales (VAS) for pain, a patellofemoral function scale (PFS), a step test and a subjective report of functional limitations. Measurements were taken on 56 PFPS patients participating in a randomized clinical trial, prior to and one month following treatment. Modest test-retest reliablility for the FIQ, VAS and step test were found. High internal consistency for the FIQ and modest internal consistency for the PFS were demonstrated. The VAS and FIQ were found to be good discriminators for measuring clinical change, while the step test was found to be reliable but poor at detecting clinical change.

Stage two, content validation, analysed 34 clinicians' ratings for components of the PFPS evaluation tool including questionnaire items and 21 clinical tests. Based on these results questionnaire items were modified and five clinical tests were selected. A clinical evaluation form, operational definitions for tests and a video were developed and piloted in stage three. Four physical therapists and three subjects (two with PFPS and one without PFPS) were evaluated using the tool.

In stage four, the PFPS tool was used to assess 41 patients and 28 age matched non-PFPS subjects. The results suggested that pain and function questionnaire components were valid measures and could be combined, as they represented one dimension of PFPS evaluation. Activity questions represented a separate dimension and were found to be primarily useful for categorization, as they did not demonstrate properties of discriminative validity and were not sensitive to clinical change. Of the five clinical tests, lower extremity alignment and flexiblity of tensor fasciae latae and rectus femoris muscle groups demonstrated reasonable levels of reliability, discriminative validity and sensitivity. The patellar orientation test was found to be limited in test-retest reliability, but appeared to have potential in regards to discriminative properties. Patellar mobility was not found to be a useful measure, as it was not able to discriminate between the patient and non-patient groups.

The results provide direction for future investigation and continued development of PFPS evaluation methods. Based on the findings it would appear that a cumulative index using the pain and function questions and the four clinical tests, excluding patellar mobility, would be a useful measure for evaluative research. However, each clinical test should be analysed when monitoring patients and determining specific treatment goals and programs.

ACKNOWLEDGEMENTS

Thankyou to my co-supervisors Dr. Art Quinney and Dr. David Magee for the direction and encouragement given to me during my program. Also, I would like to recognize the contributions of professional colleagues in the fields of physical therapy and medicine, from Saskatoon and across Canada, who participated in various stages of the research.

Finally I would like to thank Professor V.J.Cottrell, Past-Director of the School of Physical Therapy, University of Saskatchewan, for his efforts to support my research and scholarly activities.

TABLE OF CONTENTS

	Page
CHAPTER ONE	1
Clinical Measurement Tools	1
Introduction	1
Subjective Health Rating and Clinical Outcome Measure	2
Development	3
Reliability and Validity in Health Measurement	6
Reliability	6
Types of Reliability	6
Statistical Analysis of Reliability	8
Validity	10
Face Validity	10
Criterion Validity	11
Content Validity	11
Construct Validity	13
Statistical Analysis of Validity	13
Sensitivity or Responsiveness	18
Pain Measurement	21
Scaling Methods	24
Lower Extremity Impairment and Disability Measurement	28
Activity Patterns	30
Patellofemoral Pain Syndrome	32
Etiology	32
Clinical Evaluation and Diagnosis	35
Treatment	40
Measuring Clinical Change	42
Objectives and Stages of the Study	44
Bibliography	47

	Page
CHAPTER TWO	61
Paper One: Analysis of Outcome Measures used in the Study of Patellofemoral Pain Sydrome	
Introduction	62
Methodology	65
Data Analysis	68
Consistency	68
Measure of Clinical Change	68
Results	68
Discussion	71
Conclusions	75
Figures	76
Tables	80
Bibliography	88
CHAPTER THREE	91
Paper Two: Development of a Clinical Tool and Patient Questionnaire for Evaluation of Patellofemoral Pain Syndrome Patients	
Introduction	92
Review of Literature	93
Methods	97
Evaluative Study of PFPS Patients	97
Content Validation Questionnaire	97
Expert Reviewers	98
Results	98
PFPS Evaluation Tool	100
Discussion	101
Pain	102
Function	103

	Page
Activity	104
Clinical Tests	105
Conclusions	106
Figures	107
Tables	113
Bibliography	116
CHAPTER FOUR	120
Paper Three: Validation of a Clinical Evaluation Tool for Patellofemoral Pain Syndrome	
Introduction	121
Methodology	122
Data Analysis	126
Results and Discussion	129
Conclusions	139
Tables	142
Bibliography	153
CHAPTER FIVE	155
Discussion and Conclusions	155
Consistency	156
Test-Retest	156
Internal Consistency	157
Known Group Differences	158
Correlation Analysis	159
Sensitivity and Specificity	159
Bibliography	163

APPENDICIES

		Page
Appendix A-1	Occupational Rating Scale By Work Title	164
Appendix A-2	Work Activity Rating and Level	164
Appendix A-3	Sport Activity and Function Form	166
Appendix B-1	Cincinnati Knee Rating System	168
Appendix B-2	Patellofemoral Function Scale	170
Appendix C	Functional Index Questionnaire	174
Appendix D	Content Validation Questionnaire for Expert Reviewers	176
Appendix E	Clinical Evalulation Forms Initial and Final Evaluations	191
Appendix F	Patient Questionnaires Initial and Final Questionnaires	196
Appendix G	Operational Manual	205

LIST OF TABLES

		Page
Table I-1	Demographics of subjects.	80
Table I-2	Descriptive statistics for VAS Pain, FIQ and Step Test scores pre-treatment and post-treatment.	81
Table I-3	Intra-class correlation coefficients for Test-Retest of VAS and Step Test.	82
Table I-4	Spearman Rho correlation coefficient values for the total FIQ and individual questions over 3 days prior to treatment.	83
Table I-5	Descriptive statistics for Pateliofemoral (PF) Function Scale and self-report of functional limitations.	84
Table I-6	Internal consistency of FIQ using Cronbach's Alpha Model.	85
Table I-7	Reliability analysis of Patellofemoral (PF) Function Scale.	86
Table I-8	Size and direction of change for VAS Pain, FIQ, Patellofemoral Function Scale (PFS) and Step Test using mean scores.	87
Table II-1	Demographics of clinical reviewers.	113
Table II-2	VAS rating of importance of each area of PFPS evaluation.	114
Table II-3	Ranking of tests to detect clinical change in PFPS patients.	115
Table III-1	Inclusion and exclusion criteria for PFPS subjects.	142
Table III-2	Demographic data for PFPS and Non-PFPS groups.	143
Table III-3	Clinician rating of eight variables using visual analogue scales for matched PFPS and Non-PFPS groups.	144
Table III-4A	Visual analogue scores for questionnaire components	

	for matched PFPS and Non-PFPS groups.	145
Table III-4B	Activity ratings for Lipid Research Clinic Questionnaire (LRCQ) and Activity Pattern Questions for matched PFPS and Non-PFPS groups.	146
Table III-5A	Internal consistency of five clinical tests and questionnaire components for pain, function and activity using Cronbach's Alpha Model for 41 PFPS subjects.	147
Table III-5B	Intra-class correlation coefficients (ICC) for five clinical tests and questionnaire components of pain, function and activity on 28 Non-PFPS subjects.	148
Table III-6A	Correlation matrix for clinician ratings (Method 1) and subject ratings (Method 2) of pain, function and activity in PFPS group (n=41).	149
Table III-6B	Principal components factor analysis with varimax rotation of clinician ratings and subject ratings for pain, function and activity.	149
Table III-7	Diagnostic discrimination of musculo-skeletal lower extremity tests for matched PFPS and Non-PFPS groups.	150
Table III-8	Effect sizes for clinician rated tests and questionnaire components for PFPS and Non-PFPS groups.	151
Table III-9	Cumulative scores and effect sizes for two test periods for PFPS and Non-PFPS groups.	152

LIST OF FIGURES

		Page
Figure I-1	Patellofemoral Function Scale.	76
Figure I-2	Visual Analogue Scale for Pain.	77
Figure I-3	Functional Index Questionnaire (FIQ).	78
Figure I-4	Scale for rating functional limitations and change in condition.	79
Figure II-1	Activity Pattern Questionnaire.	107
Figure II-2	Content Validation Questionnaire overview.	108
Figure II-3	Initial Clinical Evaluation Form.	109
Figure II-4	Modified Lipid Research Clinic Questionnaire (LRCQ).	111
Figure II-5	Patellar Orientation Operational Definition.	112

CHAPTER ONE

Analysing the literature related to functional outcome measurement for patellofemoral pain syndrome encompasses several areas. The literature on patellofemoral conditions is extensive, as is information on health measurement. Therefore, selected studies relating specifically to this study have been reviewed in this chapter.

CLINICAL MEASUREMENT TOOLS

Introduction

As the health status of the general population has changed over the years, there has been a growing need for health indicators that reflect problems other than mortality or severe handicap or disability. Measurement of health can be based on variables gathered from objective methods such as diagnostic or laboratory tests, or from subjective methods, which rely on the subject or clinician to make a judgement regarding a particular variable. "Measurement is the process of applying a standard scale to a variable, however there is no standard scale for health. Therefore measurement relies on a number of health indicators, each of which represents an element of the overall concept." (McDowell & Newell, 1987, pp.12)

Historically health indicators have ranged from simple, nominal scales, identifying presence or absence of disease, to more complex measures that attempt to quantify health (Larson, 1991). Although some instruments have been developed for general use, most effort has gone into the development of tools to measure single disease processes. For example, the area of rheumatology has traditionally been concerned with measurement of function. The Arthritis Impact Scale (AIMS)(Meenan ,Gertman & Mason,1980) and the Health Assessment Questionnaire (HAQ, Stanford University) (Liang, Fossell & Larson, 1990; Fries, Spitz & Young, 1982; Chambers, MacDonald, Tugwell et al., 1982) are two instruments designed specifically for rheumatoid arthritis patients. More generalized instruments, such as the Sickness Impact Profile (SIP) Begner, Bobbitt, Pollar, et al., 1976; Deyo, 1988) and the Index of Well-Being (Liang et al., 1990;

Kaplan, Bush & Berry, 1976) are often modified to use with specific diseases. Although many tools stand up to statistical analyses, some of the instruments are not practical for clinical environments and may not evaluate the most relevant outcomes. To date, no tool has been developed that is appropriate for functional measurement of the general population.

A clinical environment imposes restrictions on the methods that can be appropriately and effectively used to evaluate health status. In addition to limitations in time, equipment, facility and personnel, the controlled health care setting does not simulate the real life environment in which people must function. Clinical measurement methods have been developed primarily to evaluate levels of impairment, rather than disability. The relationship of impairment to disability is not known in many cases. Therefore, the clinical tools, although reliable and valid for measuring impairment levels, may not be useful in predicting function or measuring outcomes of treatment that may affect an individual's activities of daily living.

Subjective Health Ratings and Clinical Outcome Measures

McDowell and Newell (1987) group subjective health measurements into three categories: general well-being, symptoms of illness and function. In the past, subjective health ratings have often been considered too biased and less reliable. Jette (1989) defined a subjective outcome as an observed entity which results from the subjective state of a person, based on the person's feeling; or an entity which is perceptible only to the person being assessed. In comparison, a subjective test has been defined as a measure that is determined or influenced by the beliefs or feelings of the examiner or extraneous characteristics of the subject. Subjective measurements amplify information by describing the quality of life versus merely the presence of morbidity or mortality. Insight into factors such as pain and function can never be inferred solely from physical examinations or laboratory measurements. Although the reliability of subjective data has been considered susceptible to bias, the bias can be controlled with proper study design and instrumentation (McDowell & Newell, 1987).

Arguments for considering subjective judgements as a valid approach to measurement have been derived from the field of pychophysics. Theories leading to the scaling of sensations and determining the magnitude of changes in sensations have evolved over the years. The concept of a geometric increase in brain activity, accompanied by an arithmetic increase in conscious sensation following a stimuli, was accepted for many years. However, it was found that a logarithmic relationship did not fit all types of stimuli. In the 1950s Steven's power law suggested that the response varied depending on the type of stimuli. Theories, such as these, are the basis for many of the indicators currently used to measure subjective variables such as pain (McDowell & Newell, 1987).

As Delitto (1989) suggested, subjective measures are used extensively in clinical decision making, however, there is a need for better quantification of subjective phenomena. Clinimetrics, as described by Feinstein (1987), involves the quantification of clinical data and considers five features as important to development of appropriate measurements. The features of a practical clinimetric measure, discussed in detail by Feinstein (1987), and summarized in a publication by Delitto (1989), included the following. The measure must demonstrate face validity, content validity, ease of usage and suitability for the clinical environment intended. As well, the measure must be replicable and responsive to change. Delitto(1989) pointed out that although functional assessment indices exist, they are not widely used in the clinical setting. Delitto suggested that the reasons for non-acceptance of these instruments into the clinical environment may include: reluctance on the part of the clinician to give worth to subjective data, despite experimental validation, and limited validity of the tools, especially in regards to practicality and interpretation of measurements.

Development

Determining the purpose of the tool is the first stage in development of health measurement systems. For example, in the development of functional status measurements, clinical endpoints are often combined to develop a single outcome score (Smythe, Helewa, & Goldsmith, 1982). When a battery of tests are

used together this is often referred to as a measurement index. A discriminative index is described as discriminating between subjects or groups on a dimension, when no external criterion or "gold" standard is available. A predictive index is used to classify individuals into a set of pre-defined measurement categories when a "gold" standard is not available. This is often used as a screening or diagnostic tool. An evaluative index is used to measure the magnitude of change over time in an individual or group (Kirschner & Guyatt, 1985). Different instruments or combinations of methods are required to measure each of these categories, and requirements for maximizing one index may influence or even impede the others (Wonca Classification Committee, 1990). There are methodological and statistical issues relative to the combination of single measures into indices. Information may be lost by combining results into one score. Also, the combination of several clinical endpoint measurements may not reflect functional status. For example, pulmonary rehabilitation programs have been shown to increase exercise capacity in patients, without changes in pulmonary function tests (Deyo & Patrick, 1989; Kaplan, et al., 1976). It is essential that the goals of measurement are clearly delineated prior to developing or using a tool or an index. Also, an attempt must be made to determine the importance of individual clinical endpoints relative to functional status.

Understanding and applying methods for judging the usefulness of an instrument, once a goal has been established, involves reviewing the literature and seeking evidence to support the reliability and validity of the tool. In some cases, an existing instrument may be able to serve the intended purpose of the researcher. Many scales and questionnaires are based on other instruments that may have been used clinically. However, validation studies may be necessary prior to implementing these measurement systems, if this type of evaluation has not been done.

A scale or questionnaire should be comprised of items which have been shown empirically to be characteristic of a group. There are several approaches for developing the items including: clinical observation, theory, research or expert

opinion. Clinical observation is the starting point for development of measurement tools. In fact, most research questions are developed because clinicians have observed a particular phenomena in their practice. Although this method is useful, there may be limitations related to the sample of patients or the type of practice from which the clinician is drawing their information.

In addition to clinical observation, a theory or model can assist the researcher in developing items. Theory has also been shown to have limitations, as it may emerge as inadequate or inaccurate, when tested empirically. Melzack's McGill Pain Questionnaire is an example of an instrument developed based on Melzack's theory of pain (Melzack, 1983). By using a theoretical approach, one may be able to understand and explain a patient's condition. The theoretical approach enhances the understanding of a problem and contributes to an improved body of knowledge.

McDowell and Newell (1987) described that an empirical approach to instrument development is commonly used when the measurement has a practical purpose. As an example of an empirical approach, Leavitt's Back Pain Classification Scale was developed to distinguish the source of pain as organic in origin or related to emotional disturbances. Although the scale has been shown to be a useful discriminatory tool, there is no clear explanation for how the method works (Leavitt, 1983). With the empirical approach, only the findings are presented, with limited rationale being provided for the results.

Finally, inclusion of expert opinion is a method used in the development of health measurement systems. The process for gathering opinions ranges from simply asking a colleague to review and comment on a draft of a questionnaire, to hosting a conference of recognized experts in the field, with rules governing the voting process (Streiner & Norman, 1989). The disadvantage of expert panels is that they may not reflect a range of opinion, thus the results may be skewed. This should be considered when determining the profile and requirements of expert reviewers.

There are several key points related to the composition of questions that

must also be addressed. Items must not be ambiguous, difficult to understand or use jargon. Streiner and Norman (1989) suggested that scales should not require reading skills beyond that of a 12 year old level (pp.39). This can be easily adhered to, as many computers are equipped with software programs that can evaluate the reading level of documents. Holden, Fekken and Jackson (1985) found that items with 10 to 20 characters had four times the validity coefficients of items with 70 to 80 characters. Therefore items should be as short as possible. Other problems in item construction that have been identified include: items that ask more than one question; use of terms that may prejudice the respondent; and negatively worded items, which tend to have lower validity coefficients (Holden, et al., 1985).

Once items for questionnaires or specific clinical tests have been generated using the various approaches, the next step is to ensure that the items adequately cover the domain of interest. Content validation is the term used to describe this process and it will be discussed later in the review.

RELIABILITY AND VALIDITY IN HEALTH MEASUREMENT

Reliability

Many types of error occur in measurement including bias and random errors (Wittenborn,1972). Reliability is concerned with random errors; bias is assessed by validity testing (McDowell & Newell, 1987). Internal consistency, test-retest and interrater tests are types of reliability frequently reported for functional measures. Reliability considers the extent to which a measuring instrument yields the same results on independent repeated trials under the same condition (Guyatt, Walter & Norman, 1987). Reliability is essential for instruments designed for discriminative research. In evaluative research the level of reliability required may not be as stringent, depending on the goals of the study and the experimental design.

Types of Reliability

Test-retest reliability is analyzed when information regarding the stability of a tool over time is required. Ideally one would expect that a tool should

consistently report similar scores when the test is repeated under the same test conditions. Factors contributing to error in test-retest situations include changes in the method of measurement, effects of practice or learning, and instability of the variable being measured (Helmstader, 1964; Nunnally, 1967). Pain is an example of a variable which can greatly fluctuate over time.

The amount of time between tests has been shown to influence the stability of a measure. A subject's response may be affected by recall if testing is repeated within a short time interval, as a result the assessments may not be considered independent (McDowell & Newell, 1987). If subject bias is an issue, then internal consistency can be evaluated. This form of reliability requires that two equivalent measurements are applied at the same session and focusses on inconsistency of the respondent's rating. Statistical tests used to evaluate internal consistency include KR-20 and Cronbach's coefficient alpha. These tests determine the correlation between the different versions of the measurement, thus estimating the repeatability of the test. A major limitation to this method is finding tests that measure the same variable or set of variables (Aday, 1989; Nunally, 1967).

Interrater reliability is concerned with consistency of scoring between different clinicians or raters using the same tool and protocol. Determining the error due to clinician differences provides insight into the usefulness of a tool that may be used in settings in which more than one person evaluates a subject. In most cases, a tool which demonstrates high interrater reliability will demonstrate high intrarater reliability. An individual will agree with themselves more than with someone else (Riddle, Finucane, Rothstein & Walker, 1989; Nunnally, 1967; Kerlinger, 1973).

Another factor that influences stability of measurement in clinical evaluation is the number of questions or tests included in an instrument. The length of a test is commonly restricted due to cost implications, practical considerations in the clinical environment and compliance issues. However, if it is assumed that every response has a degree of measurement error, then by averaging or summing responses over a series of questions, the error can be reduced (Norman &

Streiner, 1989, pp.8). By reducing the error, reliability of the measure will improve. Therefore, the length of an instrument must be based on a balance of the practical and statistical issues.

It is important not to assume that instruments maintain reliability when different populations are investigated (Guyatt et al, 1987; Bostram, Harms-Ringdahl & Hordemar, 1991). For example, speed of gait may be fairly consistent in an active below knee amputee population, however the same test may not be reliable with a population of rheumatoid arthritic patients, especially during an exacerbation of the disease. Validity can also be greatly influenced when different populations are evaluated.

Statistical Analysis of Reliability

A reliability coefficient (r) is generally considered as a measure of an instrument's ability to discriminate among individuals (Guyatt & Kirshner, 1985). Statistically this coefficient is defined as the ratio of the variance attributable to true differences among subjects, to the total variance. The total variance includes the true variance and the variance due to random measurement errors, which are assumed to be independent of the measurements themselves (Guyatt et al., 1987). A reliability coefficient of 1 or -1 suggests that the measurement method has no random error, therefore demonstrating that the tool can effectively measure the variable of interest. An increase in random error is represented by a reliability coefficient that moves towards 0. The result is a decrease in the ability to measure the true score. A correlation coefficient is reported with two digits following the decimal point, and a positive or negative sign preceding the digit. The magnitude of the digits represents the strength of the relationship and the sign indicates the direction of the relationship.

A correlation coefficient between variables of 0.8 is generally described as a fairly strong relationship, a correlation of 0.6 denotes a moderate relationship and a correlation of 0.2 indicates only weak association between test scores (Malgady & Krebs, 1986). The probability (p) value presented in correlation analysis indicates the level of confidence or degree of certainty that some

relationship exists between the variables of interest. The p value denotes how often the calculated correlation or a larger value would be found by chance, if no real relationship existed in the sample population. Confidence interval levels improve and the p value decreases as the magnitude of the correlation increases and also as the sample size increases.

A high correlation reflects parallelism among the variables, however the means may differ widely. For a test to be perfectly reliable, means and standard deviations must be identical, in addition to a perfect correlation. If determination of agreement among variables is important, then an intraclass correlation coefficient analysis would be appropriate, as this formula takes into account mean differences among scores and lack of parallelism (Malgady and Krebs, 1986). An intraclass correlation coefficient (ICC) reflects both systematic and random differences in scores, and is recommended as the preferred method to quantitate reliability (Guyatt et al. 1987; Stratford, 1990). Rather than measuring the correlation between two sets of scores, the ICC describes concordance, which is the extent to which the repetition of the test produces the same values under the same conditions in the same individual (Guyatt, et al, 1987). Unlike correlation analysis, which only considers relationships between two sets of scores, the ICC procedure can deal with repeated measurement data. The ICC can be calculated by using the mean-square (MS) values from a repeated measures analysis of variance model considering the number of measurements (k) on each subject.

$$R = \frac{MS(subjects) - MS(error)}{MS(subjects) + (k-1)MS(error)}$$

(Stratford,1990; Streiner & Norman, 1989)

There are a number of forms of ICC that can produce different coefficients when using the same data. Therefore it is important to consider the design of the study, when selecting the appropriate analysis and when considering the applications of the reliability results (Shrout & Fleiss, 1979).

In addition, it should be noted that measurement of reliability, using the correlation coefficient, depends on the magnitude of the variance attributable to between subject differences. The study design will indicate the requirements for differentiating between subjects or within subjects over time. If the goal of a study is to determine change within subjects over time, as in most evaluative clinical trials, the tool may not require the same degree of reliability (Guyatt et al., 1987). Conventional measurement of reliability using an ICC, considering between subject variance to total variance, may be misleading if the instrument's purpose is not defined.

Validity

Validity indicates the range of inferences that are appropriate when interpreting results of a test (Cronbach, 1971). The validity of a measure is not absolute, but relative to the domain about which the measure has been developed (Kaplan, et al.,1976). For example, an indicator or measure of function is a valid measure only to the extent that it reflects or quantifies that construct.

Various forms of validity of a measurement tool may be analyzed depending on the nature and goals of the tool. Slight differences in the categories of validity are presented in the literature, however, the same areas are consistently discussed. The American Psychological Association, the American Education Research Association and the National Council on Measurement in Education have defined three basic types of validity: criterion, content and construct (American Psychological Association, 1974). Face validity is not included as a category of validity. Included under the basic three forms of validity are various sub-categories. A detailed review of all forms of validity is beyond the scope of this section, however the most pertinent issues relative to this project will be included.

Face Validity

Although the American Psychological Association does not categorize face validity, it has been defined and discussed as the weakest form of validity (Rothstein, 1985). Face validity can not be tested, as it deals with opinions of

whether a test or treatment appears to do what it is supposed to, rather than analyzing substantive theory or evidence. It can be argued that if in the clinical environment an evaluation technique or treatment method does not have face validity, that results may be affected, as the evaluator and/or patients may not make reasonable efforts in the experiment, as they do not see the relevance of the intervention or test procedure. On the other hand, there may be situations when a researcher does not want face validity, as they wish to disguise the purpose of the test in order to reduce subject bias (Payton,1988).

In most situations, analysis of more rigorous forms of validity, such as content validity, will ensure that a test demonstrates reasonable face validity (Rothstein, 1985). It must be emphasized that face validity is not an appropriate or reliable basis for inference (American Psychological Association, 1974).

Criterion Validity

Other terms used to describe criterion validity are empirical or statistical validity (Kaplan et al., 1976). Criterion validity represents the extent to which a measure corresponds to some other observation that accurately measures the same construct. If the measures are taken at the same time this is defined as concurrent validity. If the new measure predicts a criterion value that is evaluated at a later date, this is defined as predictive validity (Rothstein, 1985; Kaplan et al., 1976).

The greatest drawback to establishing criterion validity is the existence of a criterion, that is indeed a superior measure or what is referred to as a "gold standard" (Deyo & Inui, 1984). Lack of "gold standards" are extremely evident in areas that encompass many dimensions or domains, like function or well-being. In these comprehensive areas, "... accurate expression of the total concept requires a derived measure or index number - a combination of many different, fundamental, directly observed measures." (Kaplan, et al., 1976, pp.480)

Content Validity

Content validity indicates whether instruments or evaluation methods adequately represent the domains that they are supposed to measure. As Payton

(1988) describes, content validity regulates the sampling of the construct and ensures that the measurement is comprehensive and representative of the dimensions comprising the domain. Obviously not all elements of a domain can be examined, but a sample of the elements that are most representative of the domain can be included.

The stages of content validation for an instrument to measure balance in the elderly were described in a clinical study (Berg, Wood-Dauphinee, Williams & Gayton, 1989). Phase I of the process produced a large group of items to represent each of the three domains, which had been defined based on theory and research. A group of professionals and patients were interviewed in order to generate these items. The goal of Phase II was to reduce the number of items. This was accomplished by developing a questionnaire, which consisted of the previously generated items, and having a group of clinicians rate the importance of each question. The final phase of the content validation included further elimination of items and verification of acceptability of the balance scale. A combination of professional and patient evaluation was used in this stage. Patients were scored on their performance of the activities, and were also asked to rate themselves on a five point scale. Clinicians were also asked to score patient performance based on videotapes of the balance testing. Following their viewing of the videotapes the clinicians were asked to again rate the questions, and if 60% of the expert panel agreed that an item must be included, the question was included in the scale. The authors found that the balance test was simple and easy to administer. High levels of inter and intra-observer agreement were found for individual items in the scale and the score as a whole. In addition, strong internal consistency was reported. The high level of reliability provided further evidence of content validity.

Kaplan et al. (1976) demonstrated both the importance of representativeness and proportion of dimensions in a study evaluating the validity of the Index of Well-being. Validation testing showed that symptoms, however minor, significantly decreased the overall score of the Index of Well-being. The

authors pointed out that use of any scales of dysfunctions without measures of relative importance, omits a critical element of content validity and introduces substantial bias by assuming equal weights among the items.

Construct Validity

Construct validity defines what one intends to measure and involves determining: the dimensions of the construct, the domain of the dimensions both uniquely and jointly, and the expected relations of the dimensions to each other, both internally and externally (Payton, 1988; Kaplan, et al., 1976). Establishing construct validity is essential when variables can not be directly examined but only inferred from behavior. Examples of such variables include muscle strength, pain, intelligence, and function.

In order to develop appropriate tools to measure more complex variables, a sound theoretical basis must be used. All aspects of the variable must be considered and integrated with the relevant theory for each component. Constructs are not operational definitions, however they do imply properties for proposed measures. Construction of an index is an example of an attempt to close the gap between a theoretical concept and its operational measurement. Thus the index reflects the empirical properties of the construct (Kaplan et al., 1976). Two types of external evidence for construct validity are: discriminant validity which indicates that the measure does not represent a construct other than the one it is devised to measure; and convergent validity which indicates that a test is related to other measures of the same phenomenon.

Studying the construct validity of an instrument is never finished. Various designs can be used to group construct validation studies, the types of designs include: descriptive, experimental and correlation (Goodwin & Goodwin, 1991). The correlational approach is the most widely reported method in construct validation studies.

Statistical Analysis of Validity

Since a measurement cannot be valid without being reliable, discussion of validity must always consider the issue of reliability. The maximum value of validity

is equal to the square root of the reliability coefficient. The distinguishing feature between reliability and validity is that with reliability evaluation, methods or measures are dependent, whereas in analysis of validity, methods or measures are independent.

Although various statistical procedures can be used to analyse validity of an instrument, correlation analysis is the mathematical basis for many of these tests (Malgady & Krebs, 1986). Correlation coefficients indicate how well an equation can predict one variable from another (Rothstein, 1985). The linear regression equation is most commonly used to quantitatively describe the association between variables. Factor analysis is a multivariate statistical procedure that examines the relationships among a number of variables. This technique is often used in order to reduce the number of variables.

Along with descriptive statistics, three other statistics that are commonly reported in validation studies include the coefficient of variation (CV), the standard error of measurement (SEM) and the coefficient of determination (r^2). Each of these values provides information on the variance of the measurement. The CV is the standard deviation (SD) expressed as a percentage of the mean. The CV is used to estimate the percentage of variation that can be expected in a measurement because of measurement error (Payton, 1988).

The purpose of the SEM statistic is to provide a number that represents the way in which a single score will vary when a test is administered more than once. Considering a situation in which no change has occurred, it is assumed that the score should be stable. Therefore, if variability between tests exists, this variance is solely due to error and can be assessed by SEM. The SEM can also be linked to the reliability of a variable considering the following formula: SEM = SD x SQRT(1-r). As reliability or degree of association decreases the SEM increases. The Coefficient of Determination represents the average variability that can be accounted for when predicting one measure from another measure. The statistic is calculated by squaring the correlation coefficient (r). The Coefficient of Determination indicates the average error but not the expected error for any single

measurement (Payton, 1988; Streiner & Norman, 1989).

The SEM is considered to be the most useful index of reliability, as it allows each individual score to be considered with an error term. Also the SEM assists with interpretation of the measurement (Rothstein, 1985). For example, in a hypothetical study using the visual analogue scale for pain measurement on a sample of patients, a SEM of 2 points was found during initial testing. When subsequent evaluation, after a regime of treatment was carried out, a decrease in pain of 2 points was calculated. One would have to seriously consider whether a real change had occurred in this situation, or if this change could have occurred due to error.

A common problem in functional status measurement is that distribution of responses or variability is often skewed. An example of this is seen when the Katz ADL Index, which rates self-care activities, is compared on ambulatory and non-ambulatory patients. In the institutional setting, patients receive scores over the full range of the scale, however, when administered to ambulatory patients all patients receive the highest score (Wonca Classification Committee, 1990). Many statistical techniques assume normal distribution, and in many cases functional status measurements are truly skewed. This must be considered when analyzing results, especially if different populations are being compared. Revision of instruments to encompass the entire functional range of the populations is one method for dealing with this problem. As well, there are special statistical techniques that can be used when a normal distribution is not found. It must be emphasized that many of the functional instruments only evaluate lower levels of function, therefore these tools may not be appropriate in the assessment of more active populations.

As previously mentioned, correlational analyses are most commonly reported in construct validation studies, in addition to multiple regression, factor analysis and discriminant analysis. The techniques of generalizability theory have also been used in the study of construct validity. The selection of study design and statistical analysis must be based on sound theory and hypotheses testing.

Campbell and Fiske (1959) proposed the multi-trait multi-method matrix (MTMM) as a test for construct validity. The MTMM approach analyses convergent and discriminant validity and estimates the effect of method variance on validity assessment. The MTMM matrix consists of sets of intercorrelations between a minimum of two traits with at least two independent methods for measuring the traits. The underlying theory of the MTMM approach is that tests designed to measure the same trait should correlate highly with each other. In contrast, those tests measuring one construct should not correlate with tests measuring other constructs.

The following four conditions, as described by Campbell and Fiske (1959), must be considered when examining the matrix:

1. The correlation values in the validity diagonals must be significantly different than zero. The validity diagonal correlations are defined as the homo-trait, hetero-method values (same trait, different method).
2. Each correlation value in the validity diagnonal should be higher than correlation values of that trait with other traits, measured by different methods (hetero-trait, hetero-method).
3. Homo-trait, hetero-method correlation values should be higher than hetero-trait, homo-method correlations.
4. The pattern of correlation values in the hetero-trait triangles, should be similar in both the mono-method and hetero-method blocks. The hetero-trait triangles are defined as correlation values above and below the validity diagonals.

Ferketich, Figueredo & Knapp (1991) discussed three limitations to the traditional bivariate analysis of the MTMM that affect the utility of the method. The first issue concerns the difficulty in interpreting results due to the lack of a standard criteria for assessing the magnitude of the correlations. Also, consistent patterns in the hetero-trait triangles are frequently not found, leading the researcher to attempt to interpret results when only parts of the criteria are met. Linked to inconsistency in relationships and patterns is the reliability of the measures. Ferketich et al. (1991) suggest that more attention must be paid to the

impact of lack of instrument reliability on the matrix values.

A second issue is that traits and methods must be carefully selected. A discriminant trait must be similar to, or as Ferketich et al. (1991) recommended, "substantively confused" with the trait of interest. For example, to test the construct validity of an instrument proposed to measure lower extremity function, activity pattern could be considered as a reasonable discriminant variable (trait). Often traits are used that do not correlate at all or even reasonably with the variable of interest. In regards to method selection, it may be difficult to find two or more reliable methods to measure the same trait. Although Campbell and Fiske (1959) recommend that methods should be independent, Ferketich et al. (1991) suggest that many studies violate this requirement, resulting in shared method variance. Examples of two independent methods would be patient self-report and clinician observation. Short form and long form multiple choice questionnaires would not be considered as independent methods of evaluation.

Finally, the problem of not being able to analyze and separately estimate the variances due to the trait, the method and random error is discussed as a limitation to the MTMM approach. Figueredo, Ferketich & Knapp (1991) recommended that Confirmatory Factor Analysis (CFA), a statistical technique developed by Joerskog, is an approach that can identify and quantify trait and method variances. "Widaman (1985) used CFA to construct common factors representing the latent traits and methods of the MTMM approach". (Figueredo, et al., 1991; pp 381) By identifying and quantifying trait and method contributions to test scores, intercorrelations are more clearly interpreted. Figueredo, et al. (1991) pointed out that although the CFA technique accounts for trait and method common factor variance, the statistical analysis does not correct for the other limitations discussed.

Campbell and Fiske (1959) recognized that many MTMM matrices may not produce evidence of convergent or discriminative validity. The authors suggested that although convergence between two clearly independent methods is a satisfactory minimal requirement, determining that a trait is different from other

traits is more difficult. The authors recommended that careful examination of the MTMM matrix should be used to indicate to the researcher how to proceed with development of the tool. They further suggested that as the researcher refines the measurement methods and determines which concepts need better definition, that the validity coefficients obtained at different stages in the process can be interpreted in terms of gains over previous evaluation. Convergent and discriminative validity are important properties that must be considered in the validation process of a new tool.

Sensitivity or Responsiveness

There are different definitions of sensitivity used in the measurement literature. In diagnostic testing, sensitivity reflects the ability of a tool to measure the proportion of those individuals who present with positive tests when they have the disease (Wagner, Wagner & Fletcher, 1988). Sensitivity or responsiveness is also described as the ability for a measure to detect an important change in a variable. (Deyo & Inui, 1984; Deyo & Centor, 1986; Guyatt, Deyo, Charlson, Levine & Mitchell, 1989) Sensitivity, used in this context, is an essential element of evaluative research, but there is no consensus on the best approach to measuring this property (Guyatt, et al.,1987). By increasing the increments or descriptors on a scale, sensitivity or responsiveness of a measure may be improved. However, by doing this you may also decrease the reliability of the measure and decrease the specificity of the tool (Guyatt et al., 1987). If a tool lacks specificity, it does not yield normal results when there is no abnormality (Deyo, 1988; Noyes, Barber & Mangine, 1991).

The sensitivity of four different types of one-legged hop test was examined in a sample of anterior cruciate ligament (ACL) deficient knees (Noyes, Mooar, & Barber,1991). Previously these investigators had studied 35 patients and found low sensitivity rates for the functional tests. The aim of a followup study was to see if an increase in sample size would improve the sensitivity of the tests. The authors had previously studied 97 non-affected subjects and determined values for limb symmetry. In the 1991 study, 67 patients were evaluated, almost twice

as many patients as in the 1987 study. The results of the second study were similar to the first, with no improvements in the sensitivity of the functional tests found. The authors concluded that these tests should not be used for diagnostic or screening tests and were only useful for confirming a previous diagnosis.

The ability of an instrument to detect individual differences in change scores is often referred to as reliability of the change score. This is depicted as the ratio of the systematic difference between subject change scores compared to the systematic difference between subject change scores and variance error (Guyatt, et al., 1987). If an intervention resulted in a uniform response to treatment, all patients would improve an equal amount. As a result the variance of the change score would be 0 because post and pre-test scores stayed the same except for a constant. Therefore the reliability of the difference score would be 0. A perfectly uniform response to treatment would represent the ideal for the use of change scores to measure treatment effects, yet the reliability coefficient for change scores would be 0. Therefore, a reliability coefficient should not be used as an index appropriate for assessing the ability of an instrument to detect change (Deyo, 1988; Guyatt, et al., 1987). Measurement of sensitivity or responsiveness , analogous to the reliability coefficient, is required to describe an instrument's ability to detect overall treatment effect.

In a study on functional status of low back pain, work status was identified as an insensitive indicator of change (Deyo, 1988). The many factors that affected work including job satisfaction, position, role and responsibilities made it difficult to measure changes in work status. For example, someone may have stayed at work because of financial needs or motivation, despite serious low back problems that affected function. Deyo (1988) discussed the use of receiver operating characteristic curves (ROC) in measuring the responsiveness of four functional questionnaires on low back pain patients. In this study, different rankings of questionnaires were found when analysis of difference scores was compared to analysis using the ROC method. The ROC method involves plotting the true positive values (or real change) against false positive values (1 - specificity =

random or other non-specific score variability). The areas under the curves are calculated and then compared. The advantage of the ROC method of analysis is that it considers change, as well as distinguishing improvers from non-improvers. Although an instrument may be highly sensitive to change it may also be non-specific and this jeopardizes the usefulness of the tool. In another situation a tool may be specific but not sensitive.

The following three approaches for determining sensitivity or responsiveness for functional scales were commonly found in the literature. The first method was comparison of correlations for functional change scores and changes in clinical variables. This is a useful method as it considers all data and discriminates between those who get better and those who get worse. A limitation to the method is that it does not permit quantification or formal statistical comparisons of scale responsiveness (Deyo & Centor, 1986).

Pre and post treatment score comparison is presently the most common method used for assessing scale responsiveness (Guyatt, et al., 1987). Often a paired t-test is used and the largest t-value is considered most sensitive (Deyo & Centor, 1986). A paper by Liang and colleagues (1990) described this as an index of standardized response means (mean response/response standard deviation). Deyo and Centor (1986) suggested that there are problems with this method, as it does not demonstrate the ability to detect the differences between those people who improve and those who do not. Streiner and Norman (1989) suggested that change scores should only be used when the reliability of the measure exceeds 0.5. Reliability is a necessary pre-requisite for the appropriate application of change scores.

The ROC method permits formal statistical comparisons between scales and considers improvers and non-improvers (Deyo, 1988). Unfortunately this approach requires that the external outcome criterion be dichotomous, rather than preserving information about the degree of change (Deyo & Centor, 1986). Since gold standards seldom exist in functional measurement, several external criteria may be used (Deyo, 1988).

Other strategies for evaluating sensitivity have been discussed in the literature. A sensitivity coefficient is expressed as the ratio of the variance of the change compared to the variance of the change and variance error (Norman, 1989). This coefficient reflects the proportion of the variance in the change score due to true change. Although the reliability coefficient and sensitivity coefficient are related, they are not identical. Both use the same error term, but the sensitivity coefficient considers the magnitude of the variance of change scores. Reliability and sensitivity must be considered together.

Another statistical method used to analyse responsiveness is expressed as the ratio of a clinically important difference compared to the variability in stable subjects. The formula is expressed as the square root of 2 times the mean square error (MSE). If a clinically important difference is not known the mean score change with treatment is compared to the mean score change in stable subjects (Guyatt, et al., 1987).

Regardless of the method chosen to evaluate responsiveness of a tool, the investigator must consider the effect of systemic and random error on the power of the test (Kirschner & Guyatt, 1985; Tuley, Mulrow & McMahan, 1991). If the measurement error is large, then the power of the test will be low, unless large sample populations are attainable for study. If a tool only detects differences when large populations are studied or the effect size is large, the tool would be inappropriate for evaluative research in rehabilitation. It is common in the literature to find articles that report levels of reliability and validity, but fewer studies specifically address the issue of sensitivity. Without determining the responsiveness of a measure, an individual can not judge the effect of a clinical intervention accurately. In the past, some clinical trials may have incorrectly found interventions to be ineffective, when in fact the measurement tool lacked the ability to detect clinically significant changes.

PAIN MEASUREMENT

The objective assessment of pain is a challenging and critical task for health professionals. Accurate and reliable measurement of pain is necessary for

diagnostic purposes and for evaluating pain management programs. Pain is a multi-dimensional phenomenon that is commonly categorized into sensory and psychological components. Sensory refers to the anatomical, physiological and chemical factors, whereas psychological is related to the psycho-social and affective variables (Melzack, 1983). Clinically one often measures both components and further discriminates between acute and chronic types of pain, depending on the length of time that the pain has persisted. Melzack (1983) described pain lasting longer than 6 months in duration as chronic in nature. The extensive literature in the area of pain management, demonstrates the multi-faceted nature of the variable, as therapy may be directed at the physical pain stimulus or the patient's reaction to it (McDowell & Newell, 1987).

Many laboratory techniques are directed towards indirect pain measurement by focussing on the stimulus, rather than the response. Quantitative models for examining pain behavior include: traditional threshold, magnitude estimation, multi-dimensional scaling and signal detection theory (SDT) or sensory decision theory. Of these models only the SDT quantifies pain into sensory and attitudinal components. Direct measurement can be incorporated by using direct scaling or magnitude estimation of the pain intensity. A limitation to the laboratory model of pain measurement is that acute pain is investigated, not chronic pain (Melzack, 1983).

Clinical pain measurement instruments have been extensively researched. Many experimental designs have been used to evaluate the effects of interventions and to investigate the methods of measuring pain, especially those that compare attributes of patients. An important distinction that must be made in clinical pain studies, is whether the evaluation is static or dynamic in nature. An example of a static condition is one in which an instrument is compared to other instruments measuring pain behavior at one time. Also, reliability assessment of an instrument, assuming that the variable is stable over the time period, is considered a static condition. Dynamic conditions refer to studies that assess the effects of an intervention, such as analgesic medications, transcutaneous electrical

stimulation, or counselling programs on pain behavior (Gracely, 1992). Several authors have suggested that various interventions may have an affect on the pain measurement instrument (Heft, Gracely & Dubner, 1984; Melzack, 1985; Gaughan et al., 1990).

Four aspects of clinical pain measurement, as described by Fordyce (1983), include: pain behaviors, functional impairment, health care utilization and associated problems. To some extent all of these factors can be used as measures of pain. Reporting of pain is influenced by biological, social and psychological factors. Biologically there may not be a linear relationship between pain and the extent of tissue damage. Many cultural and individual factors, including sex, environment, personality and age have been shown to influence pain responses (Melzack, 1980, McDowell & Newell, 1987). Other factors shown to alter the reporting of pain include: attitude, anxiety and psychological intervention such as suggestion, hypnosis, placebo, attention and biofeedback. (Melzack, 1983) In children, demographic, medical factors and psychological factors have been shown to influence ratings of pain (Manne, et al, 1992). Compounding these variables is the fact that pain, itself, may be rather unstable, depending on the condition.

A wide range of pain measurement methods have been investigated including those which focus on the person's subjective response, concentrating primarily on the intensity, the type or temporal fluctuation of the pain sensation. The sensory discrimination theory is not widely used in the clinical setting, as simpler intensity scales are preferred. The techniques most commonly used to measure pain include: verbal records or written descriptions of pain, ratings based on observable behaviors, and analogue scales. Various questionnaires have been developed that measure the intensity and duration of pain using numerical or descriptive scales. Diaries or charts that record changes in pain and medication intake over time are also used (McDowell & Newell, 1987). A comprehensive approach to pain measurement is found in the McGill Pain Questionnaire which has been widely used in both clinical work and research.

This questionnaire evaluates 3 dimensions of pain as theorized by the Gate Control Theory (Melzack and Wall, 1965, Holroyd, et al., 1992).

Similar to other areas, controversy exists as to the best method of evaluating changes in pain measurement scores. Pre-test versus post-test scoring has been considered inappropriate because the difference will be affected by the magnitude of the baseline measure. Scott and Huskisson(1979) recommended using pain relief as the dependent variable, in order to reduce the effect of the initial score. Measurements of behavioral changes are often used as indicators of pain levels, although there is some debate as to the relationship between behaviors and pain. Recording changes in function, medication dosage, and observable body language are examples of behavioral measurements. Many other behavioral measurements have been described (Craig & Prkachin, 1983). Fordyce (1983) reported that no relationship existed between recorded activity levels and pain rating over a two week period of 150 chronic pain patients. Subjects completed a daily diary including time spent on activities, such as standing and sitting, and rated pain severity on a zero to 10 scale. In addition, scores from an Activity Pattern Indicator (API) which consists of 63 items of common daily activities, also showed no differences in performance of the activities in patients reporting high pain levels versus patients with medium and low pain levels. These findings have also been confirmed by Sanders (1980) who found a low relationship between pain levels and activity levels, reported in a diary, as well as the actual measure of activity recorded with equipment that recorded the amount of time spent moving in an upright position in chronic pain patients.

SCALING METHODS

In considering approaches to the measurement of responses, one must consider the level of measurement. Identification of nominal, ordinal, interval and ratio variables is essential for determining appropriate statistical methods. As well, the experimental design, including assumptions and limitations, and the format of the questionnaires or scales are influenced by the level of measurement.

Categorical judgements are often included in response scales, when only

two choices are possible. In some cases, categorical questions are used when the response is not categorical, but continuous in nature. Many of the health indicators of interest in evaluative research can be categorized as continuous or ordinal. McKenzie and Charlson (1986) reported on standard requirements of ordinal scales for clinical trials. Methodological problems of single state scales, which measure patients at a single point in time, and transition scales, which directly measure change in patient's status were discussed. For both types of scales the following standards were recommended. Individual ranks must be clearly defined and responses must be mutually exclusive and ordered in a hierarchial progression. Also, the scale must be able to equally detect improvement and deterioration.

There are many techniques used to measure continuous variables including: direct estimation, comparative and econometric methods (Streiner & Norman, 1989). Only direct estimation methods will be reviewed in this chapter. Direct estimation methods include: visual analogue scales, adjectival scales and specific scaling methods such as the Likert scale.

The visual analogue scale has been used for self-reporting of subjective experiences like pain intensity or functional limitations. This method can be used in evaluative research, when change in the subjective variable is important. The tool has been shown to be a simple, powerful, sensitive and reproducible instrument (Scott & Huskisson, 1979; Melzack, 1983; McDowell & Newell, 1987). Visual analogue scales have been used extensively in psychological testing and Huskisson (1974) developed the approach of using these scales for pain measurement. The scales are straight lines, commonly 10 cm. in length, presented vertically or horizontally. At each end of the scale are descriptors indicating no pain and severe pain. Huskisson's (1974) descriptors were "pain as bad as it could be" and "no pain". Adjectival scales are similar to the visual analogue scales, except that various descriptors at specific intervals are added to the line. Scott and Huskisson (1979) reported that scales were more sensitive if descriptors were only used at the extremes of the scale, as a tendency for clustering of responses

next to each descriptor has been found.

Scoring of the visual analogue scale is typically done by measuring the scale in millimeters from the lower end of the scale. In a report by Huskisson (1974), a 20-point grid, superimposed on the line was considered an acceptable alternative, as this approach represents the maximal discrimination people can use in recording pain levels. Huskisson (1982) recommended that non-parametric analyses of data should be used for scores, as results are not normally distributed and data is ordinal in nature. However, many researchers using this form of evaluation have found that parametric analysis is appropriate as assumptions are not violated (Gaito, 1982; Streiner & Norman, 1989).

Scott and Huskisson (1979) reported high correlation between successive measurements of pain severity using the VAS. The correlations between vertical and horizontal visual analogue scales have been reported as 0.89 to 0.91. When the VAS was compared to the McGill Pain Questionnaire, correlations ranged from 0.60 to 0.63. Huskisson (1974) reported 0.75 correlation between a vertical visual analogue scale and a four-point descriptive pain scale.

Downie, Leatham & Rhind (1978) also reported that the VAS used for rating pain has good correlation with other pain scales. However, these authors recommended that a 10-point numerical score may be a more appropriate method for scoring, in order to decrease the number of increments on the scale, as they felt this introduced increased measurement error. Reducing the increments on the scale will also reduce the sensitivity of this form of evaluation. Therefore, in order to reduce error, proper orientation to filling out the scale, and assistance when initially completing the form should always be performed. Some controversy exists in regards to the method for re-evaluation of pain using the scale. Scott and Huskisson (1979) reported that an important source of error in serial measurement of subjective states was raters knowledge of previous scores. In a study on 92 patients with painful rheumatoid arthritis, pain measurements with and without previous knowledge of scores were compared. The authors concluded that patients should be shown previous scores, as it leads to greater precision,

especially if measures are recorded over a long time period (Scott and Huskisson, 1979). Dixon and Bird (1981) did not concur with this method, as they suggested that change scores of pain do not correlate with other indicators of disease changes.

Flandry and colleagues (1991) analyzed subjective knee complaints using visual analogue scales on 182 patients presenting with a variety of knee conditions. Criterion validity of the scales was reported as high, when the scales were compared to other subjective evaluation methods. The authors recommended that visual analogue scales brought greater sensitivity and statistical power to data analysis, when compared to categorical methods. Budiman-Mak, Conrad & Roach (1991) also reported that visual analogue scales were effective measurement tools for determining levels of pain, disability and activity in patients with foot pain.

There are many other types of scales, especially in the area of attitude evaluation. An attitude scale consists of a series of statements that reflects the cognitive, emotional and behavioral components of attitude. The subject is given a series of statements and is asked to respond to them. Likert, Guttman and Thurstone scales are examples of attitude scales (Streiner & Norman, 1989). A series of statements, with responses framed on an agree to disagree continuum, are the basis for these scales.

When evaluating the results of scaling methods, two issues that must be considered are weighting of items and transformation of scores. Rather than simply summing the scores of all items, weighting of items may be a necessary procedure. Although a controversial issue, it appears that if a tool has fewer than 40 items, weighting may improve the validity coefficient (Nunnally, 1970). Also, if an index consists of unrelated items, then it may be worthwhile to run a multiple regression analysis, to determine empirically whether the predicting ability of the scale can be improved (Streiner and Norman, 1989). In some cases, there may be a need to transform the raw score in order to facilitate interpretation and analysis of data. If two instruments with different scales are compared a method

for standardizing the scores must be used. Four common approaches to transforming raw scores include: percentiles, standard (z) scores, standardized (t) scores and normalized scores.

LOWER EXTREMITY IMPAIRMENT AND DISABILITY MEASUREMENT

Most clinicians are familiar with evaluating impairment, which is defined as "any loss of psychological, physiological or anatomical structures or function" and "disturbances at the organ level" (World Health Organization (W.H.O.), 1980). Ratings of structural involvement and functional changes have often been arbitrarily assigned to measure levels of impairment. Measurement of ligamentous laxity is often reported according to the degree of gapping that occurs during passive testing of a joint. Assigning a grade of impairment for articular cartilage lesions at arthroscopy considers the appearance of the cartilage, the extent of involvement, the diameter of the lesion and the location of the lesion. Rating function according to the perceived difficulty in performing defined activities measures functional impairment. Although these rating systems improve the standardization of testing, validation of these measurement methods is essential if they are to be used in monitoring change and determining clinical interventions. W.H.O. (1980) defines disability as an inability for a person to carry out activities considered to be within normal ranges for a human being. Evaluation of disability encompasses more than medical impairments. Socio-economic, cultural, and environmental factors influence the resulting disability. Disability rating scales are few in number and as a result many clinicians, especially physicians determining level of disability for compensation purposes, resort to using impairment ratings alone (Noyes, Mooar, Barber, 1991).

A number of scaling systems based on patient questionnaires and clinical tests have been developed to evaluate lower extremity impairment. Many of the instruments have been designed to evaluate clinical change in patients with specific joint pathology. As the body functions as a kinematic chain, the measurement systems must reflect the impact of the joint changes to overall function. As a result, similar activities are often included in functional evaluation of

different joints conditions. One of the problems with integration of the body systems is that areas can compensate for each other, making it difficult to use function as an accurate indicator of clinical change.

Development of functional outcome measures for total hip replacement patients has been frequently discussed in the literature (Andersson, 1972). Most instruments have consistently demonstrated a difference between excellent and poor results following hip surgery, however, refinement of the tools has been indicated, in order to improve the responsiveness of these measures (Guyatt, et al.,1987). The domains of pain, walking, function, range of motion and muscle strength have been commonly included in hip measurement systems. In a questionnaire recently developed for the assessment of outcome following total hip arthroplasty, a 100-point scale, with equal weight given to the domains of pain, walking, function and overall impact of arthritis, was used (Johanson, Charlson, Szatrowski & Ranawat, 1992). Reproducibility of the results of the hip questionnaire was analyzed by comparing repetitions of the test within two weeks of the first test on a sample of clinically stable patients. Questions were removed from the initial questionnaire if the kappa statistic was less than 0.6. Validity of the tool was evaluated by comparison of the results of the hip questionnaire, the arthritis impact measurement scale and a walking test. Reasonable correlation values were reported between the different measures. The questionnaire was found to be responsive to the change in the clinical condition of the patient, although certain domains showed higher degrees of responsiveness than others.

A Foot Function Index (FFI), developed to measure the impact of foot pathology on function, demonstrated good test-retest properties ($r=0.69$ to 0,87) and evidence of construct and criterion validity in a sample of 87 patients with rheumatoid arthritis (Budiman-Mak, et al., 1991). Pain, disability and activity restriction were defined as the domains of interest. Visual analogue scales were used to rate all items under each domain. The tool was composed of 23 items divided into 3 subscales, representing each domain. Ability of the FFI to detect changes in clinical status was investigated by examining the association between

changes in the FFI and the number of painful foot joints. Painful joint count is considered to be an objective indicator of disease activity. Changes in the number of painful foot joints correlated significantly, but weakly, with changes in the activity limitation subscale (r=0.34), but did not correlate with the disability subscale scores (r=0.11). The change in the number of painful foot joints correlated moderately with the change in the total FFI scores (r=0.45) and the pain subscale scores (r=0.47). The authors recommended that this tool was a practical method for measuring change in foot function in an outpatient setting. (Budiman-Mak,et al.,1991) Testing the tool on other populations, other than rheumatoid arthritis patients, should be performed prior to generalizing its usefulness to evaluation of any condition involving foot pain.

Evaluation of lower extremity impairment resulting from knee conditions has been frequently studied, and although many of the knee rating systems have been developed primarily for patients following knee ligament reconstruction, many aspects of these systems provide the basis for evaluation of other knee conditions. Components of many of these impairment scales include ratings of: aspects of the patient's history, radiological findings, physical findings observed or tested during general physical evaluation of the joint and specific evaluation of certain structures (ie. ligaments). Function, activity patterns considering occupation, activities of daily living and sports, articular cartilage involvement, symptoms, and rehabilitation compliance are also frequently included (Noyes, 1990).

ACTIVITY PATTERNS

Methods for evaluating activity patterns must cover many aspects of daily living including: occupational, sports and recreation and normal day to day functions. Occupational limitations have been examined most frequently by scales that categorize occupations based on factors including: intensity of work, weight bearing status, and type or nature of work (Grandjean, 1982). Granger and McNamara (1984) described six categories of work intensity: homemaker, student, competitive worker, non-competitive worker, retired person due to age, and too disabled to work. Employment status of full-time, part-time, adjusted load and not

working were also rated. Tegner & Lysholm (1985) combined work and sport activities to provide a total activity rating, with six levels of work described ranging from total disability to heavy labour. Noyes et al. (1990) compared two different types of occupational rating scales in a prospective study of 50 patients with knee pathology. One scale rated the intensity of occupations by work title using a gradient ranging from zero to ten (Appendix A-1). The second scale considered specific work activities and rated the intensity, frequency and duration of tasks (Appendix A-2). Less than 50% of the patients rated themselves equally on each scale and during a follow-up interview 75% of the patients reported difficulty selecting a work activity based on work title only. The authors recommended that the activity-related scale had potential use for measuring work-related limitations. However, validation of the instrument is required, as this study was only the first stage of development (Noyes, Mooar, Barber, 1990).

The assessment of return to sports activities and a rating of sports activities vary among authors. Different numerical rating systems have been defined for determining levels of sport. Ten categories including work activities were described by Tegner and Lysholm (1985), and five gradients from very strenuous to light, separated into competitive and recreational activities, were developed by Straub and Hunter (1989). The results of treatment are commonly judged by the percentage of patients who return to sports or recreational activities following an injury. In order to compare treatment results, study populations must be similar. The intensity and frequency of sports participation and activities of daily living (ADL) must be analyzed, in addition to the identification of extraneous variables that influence return to activity.

Noyes et al. (1989) identified five parameters related to activity levels, as major sources of error in existing rating systems, and proposed a format for analysis of sports participation and ADL (Appendix A-3). This participation scale considered the frequency of play and the knee functions performed with activities sub-classified into three categories of knee function including: jumping, hard pivoting and cutting; running, twisting and turning; and no running, twisting or

jumping. Evaluation of change in sports activities, self-assessment of ADL and sport involvement, and problems with overall sports participation were also integrated into this system. Another weakness in rating systems, discussed by Noyes et al. (1989), was the failure to detect patients who continue to participate in activities, despite limitations or pain. Study methodology must address this issue, and prevent these people from inappropriately being reported as successful outcomes.

PATELLOFEMORAL PAIN SYNDROME

Etiology

PFPS is a condition that is characterized by pain associated with activities that require movement of the patellofemoral joint during lower extremity activities. The condition is commonly reported in adolescent populations and young adults, with an insidious onset of vague aching pain over the peripatellar region. Often the pain is bilateral and aggravated by activity, especially stairs, squatting and prolonged sitting (Reid, 1992; Goldberg, 1991). Medical terminology for this condition varies with the following descriptions commonly used: patellagia, anterior knee pain, peripatellar syndrome, malalignment syndrome, patellar arthralgia and chondromalacia. The literature on diagnosis and management of patellofemoral conditions is extensive. The prevalence of this condition in the adolescent population has been reported as high as 30% (Fairbank, Pynsent, van Poortvliet & Phillips, 1984), while in the athletic population 25% of participants may complain of this condition (McConnell, 1986).

The exact etiology of the condition has not been determined and it is quite apparent that more than one cause is responsible for this condition. Eisele (1991) reported that the most common etiologies included: patellar tracking disorders and instability; tendinitis; bursitis; synovitis; and traction apophysitis. In addition to the status of the articular cartilage, other factors that are frequently considered in classification of PFPS are: patellar tracking, proximal and distal biomechanical relationships, trauma, and local vasculature (Reid, 1992; Wilson, 1990; Bourne, Hazel, Scott, et al., 1988; Paulos, Rusche, Johnson & Noyes, 1980; Ficat &

Hungerford, 1977). Abnormal tracking appears to be one of the underlying causes of PFPS, frequently discussed (Insall, 1979). Changes in muscle balance, ligamentous and soft tissue stabilizers, or the osseous structures may result in the abnormal distribution of forces at the patellofemoral joint. As articular cartilage does not have nerve endings the source of pain is unclear. Pain theories presented suggest that the pain sources may include: subchondral bone, synovium, capsule and venous engorgement (Minns, Birnie & Abernathy, 1979; Abernathy, Townsend, Rose & Radin, 1978; Waisbrod & Treiman, 1980).

Reid (1992) described a diagnostic classification for anterior knee pain that considered signs and symptoms as well as physical and radiological findings. The following categories were defined: anterior knee pain, odd facet syndrome, alignment syndromes, lateral hyperpressure syndrome, plical syndrome, patella alta or baja, osteochondral injury, overuse syndromes of tendon, overuse of trauma to bursa, fat pad syndromes, reflex sympathetic dystrophy, osteoarthrosis and intra-articular mechanical (Reid, 1992; pp 366).

Damage to articular cartilage due to excessive loading has been discussed by many authors (Reid, 1992; Noyes & Stabler, 1989; Goodfellow, Hungerford & Zindel, 1976). With changes in the organization of the collagen fibers of the articular cartilage, energy absorption is decreased leading to increased compressive loads on the subchondral bone. Nociceptive fibers present in the subchondral bone respond to this excessive loading by producing pain. Destruction of the cartilage can progress to osteoarthritic joint changes. This type of PFPS is more commonly referred to as chondromalacia and requires arthroscopic diagnosis. Several systems have been developed to grade chondromalacia depending on the degree and nature of articular cartilage damage (Noyes & Stabler, 1989; Outerbridge, 1975; Metcalf, 1982). Leslie and Bentley (1978) found that 49% of PFPS patients did not demonstrate pathological changes in articular cartilage at arthroscopy. Other authors have questioned whether articular cartilage damage is directly related to the presenting symptomotology (Devereaux & Lachmann, 1984). The reasons why excessive loading occurs in

some people and not others is not known. Predisposing factors that may contribute to this pathology include: trauma to the joint, tightness of soft tissue and muscular structures such as the lateral retinaculum, and biomechanical changes in other joints in the kinematic chain.

The shape of the patella and femur, the relationship, power and endurance of muscle structures surrounding the knee, the soft tissue properties of the retinaculum, and the anatomical alignment of the lower extremity directly influence the tracking of the patella. The biomechanical functions of the patellofemoral joint are to assist knee extension by increasing the lever arm for the quadriceps muscle tendon and to distribute the compressive stress on the femur by increasing the surface area (Nordin & Frankel, 1989). During knee movements the patella glides along the femoral articulation. With the knee fully flexed the patella is securely situated in the intercondylar groove and as the knee extends the patella moves out of the groove with decreased surface contact area resulting. Three basic forces control patellar tracking. The medial force is primarily produced by the oblique head of vastus medialis, while the lateral patellofemoral ligaments create the lateral force and the patellar tendon the distal force (Henry, Goletz & Williamson 1986). Patellofemoral joint reaction (PFJR) forces vary depending on the type of activity. Goodfellow et al. (1976) described the normal patellofemoral contact areas during varying degrees of knee flexion. It has been theorized that contact areas and pressures are related to symptomotology. The source of pain, as previously mentioned, remains unclear.

The magnitude of joint reaction forces increases with the amount of knee flexion. Patellofemoral joint forces of 0.5 times body weight during normal gait and more than three times body weight while climbing and descending stairs or squatting have been reported (Fulkerson & Hungerford,1990). Patellofemoral joint reaction forces while descending stairs have been reported to be two times greater than during stair climbing. The difference in joint forces is attributed to the greater quadriceps muscle force and the increased knee flexion during stair descent (Nordin & Frankel, 1989). Open kinematic chain exercises, like knee

extension with resistance, have been shown to produce PFJR forces in excess of stairs and squatting activities (Fulkerson & Hungerford, 1990).

Patellar balance can be affected by many factors including the angle at which the patella articulates with the femur. This angle known as the Q-angle represents the position of the patella relative to the pull of the quadriceps muscle group. Although many clinicians routinely evaluate the Q-angle, there is much controversy as to the validity of the measure. Some authors have suggested that the angle can vary significantly in non-symptomatic people and can present as normal in PFPS individuals. More recently the A-angle has been proposed as another measure of patellar position. This angle represents the position of the patella relative to the tibial tubercle. In a recent study with 30 PFPS patients and 30 control subjects , the A-angle was found to be reliable when taken by the same examiner. Also, this study reported a significant difference in the A-angle between the patient and control subjects (DiVeta & Vogelback, 1992).

Clinical Evaluation and Diagnosis

The patient with PFPS is typically an adolescent complaining of an insidious onset of pain or aching over the peripatellar region, often bilateral, aggravated by stair climbing and prolonged sitting (Reid, 1992; Goldberg, 1991). Significant swelling and locking rarely occur, but a sensation of giving way, cracking or grating are often reported. Popliteal discomfort is occasionally reported in more chronic cases (Goldberg, 1991).

As patellofemoral pain can be referred from other regions, other diagnostic entities must be considered. Tumours, ligamentous and meniscal involvement, tendinitis, slipped capital femoral epiphysis (Goldberg, 1991), Legg-Calve-Perthes disease, bursitis, subluxation or dislocation of the patella, osteochondritis dissecans, and fat pad impingement are some of the common diagnoses associated with similar symptomotology. A lower quadrant scan should be performed routinely on these patients as pain may be referred from the spinal region. This type of examination screens patients for potential lumbo-sacral, pelvic, and hip dysfunctions, and is the initial step in determining postural change

including lower extremity malalignment.

Evaluation of ligamentous stability should be performed to rule out knee pain due to tibio-femoral instability. Ligament testing has been extensively researched. Most commonly the Lachman and Anterior Drawer tests are performed to evaluate the anterior cruciate ligament, while the Posterior Drawer test is performed to evaluate the posterior cruciate ligament. The lateral and medial collateral ligaments are commonly evaluated by performing varus and valgus stress tests of the knee at 0 and 30 degrees of knee flexion. In addition to ligamentous evaluation, clinicians routinely perform a meniscal test, such as McMurrays test, to challenge the integrity of the medial and lateral menisci.

A thorough physical examination is necessary in the evaluation of patellofemoral pain syndrome, and should include inspection, movement, palpation and special testing of the lower extremity and the patellofemoral structures. An excellent review of clinical evaluation is reported by Reid (1992). Lower extremity alignment is observed in standing and supine, to detect possible biomechanical abnormalities of the patellofemoral joint from anatomical changes such as femoral anteversion, genu valgum or varum, tibial torsion and pronation of the sub-talar and trans-tarsal joints (Reid, 1992; Goldberg, 1991; McConnell, 1986; Bourne, et al., 1988). Measurement of Q angle is frequently cited in the literature. The normal average Q angle is considered to be between 13 - 15 degrees with the outer limits of normal considered to be 20 degrees (Ficat & Hungerford, 1977; Gruber, 1979; Insall, 1979; Malek & Mangine, 1981). Although these values are commonly used in diagnosis, caution must be exercised in interpretation, as large population studies have not been used to determine these values. Similar to postural evaluation, large variations in the normal angles in the asymptomatic population may be found, if large population investigations were performed.

The apprehension test and manual mobilizations of the patella have been used to determine patellar stability (Reid, 1992; Kolowich, Paulos, Rosenberg, 1991). Normal values for patellar movements have not been reported and no literature was found reporting the reliability or validity of apprehension testing.

Reproduction of pain with patellofemoral joint compression, palpation of medial and lateral facets of the patella and contraction of quadriceps with external manual resistance to superior movement of the patella (Clarke's sign) have been commonly considered positive indicators of PFPS (Chesworth, Culham, Tata & Peat, 1989; Insall, 1979; Gruber, 1979;). Patellar crepitus is commonly reported by PFPS patients, but has also been found in non-symptomatic individuals. Effusion is not commonly found with the typical PFPS patient, however in acute cases swelling can be present (Reid, 1992). Tightness and/or tenderness to palpation of the medial and lateral retinaculum, quadriceps and patellar tendons, bursae and fat pads are frequently reported (Reid, 1992; Goldberg, 1991).

Quadriceps atrophy and decrease in muscle tone have frequently been associated with patellar instability and mal-tracking (McConnell, 1986; LeVeau & Rogers, 1980). Although clinical methods for determining limb girth have been established, validity of measurements are questionable. Doxey, 1987 assessed limb girth in patients with patellofemoral pain syndrome. This authors found that girth measurements can not detect muscle atrophy in subjects with leg injury. Reasons for the poor sensitivity of girth measurements to detect significant change may be due to increased fat content within the thigh, thus negating a girth decrease. Young, Hughes, Round & Edwards (1980) reported that quadriceps atrophy is usually much greater than estimates with limb girth measurements.

Muscle performance has frequently been evaluated in patients with PFPS (McIntyre & Robertson, 1992; Hantten & Schulthies, 1990; Souza & Gross, 1990; Voight & Wieder, 1991; LeVeau & Rogers, 1980; Milgrom et al., 1991; Bennett & Stauber, 1986). Much of the research has concentrated on the theory that a neurophysiologic motor control imbalance between vastus lateralis (VL) and vastus medialis obliquus (VMO) muscles is present in PFPS (Souza & Gross, 1990; Voight & Wider, 1991). The biomechanical changes that result from this imbalance have been linked to the knee symptoms. Stair climbing has been evaluated in PFPS patients, with muscle activation patterns most frequently analyzed. VMO and VL activation patterns have been shown to be greater in ascending and

descending stairs than during isometric muscle contraction during knee extension. Souza and Gross (1991) reported that PFPS patients may have abnormal VMO to VL activation patterns during isotonic activities, like stair climbing. This pattern may not be evident during isometric types of exercise.

Bennet and Stauber (1986) compared muscle torque values in non-affected subjects and a sample of PFPS patients. Decreased eccentric torque values for the knee extensors were found with the patient population. Concentric torque values were only decreased between 30 to 60 degrees of knee flexion. Irregularity in torque patterns have also been reported during knee extension testing (Hoke, Howell, & Stack, 1983). Deficits in hamstring muscle torque have also been reported in athletic populations with PFPS during isokinetic testing (Kibler, 1987). Interpretation of muscle evaluation results, such as isokinetic testing, must be made cautiously. Often patients may have symptoms bilaterally, which makes comparison between sides unhelpful. Muscle evaluation is useful in order to monitor changes during rehabilitation programs, but clinicians must not base their programs solely on the results of these tests. The relationship between muscle performance variables, such as peak torque, and function is not clear. How much change has to occur in muscle parameters prior to complaints of symptoms or functional changes, is not known. Further work in this area is needed.

Although clinical methods for evaluating flexibility of the tensor fasciae latae, hip flexors, hamstrings and gastrocnemius-soleus muscle groups have been developed, normal values vary greatly and are highly subjective. The following are examples of tests and parameters that are frequently used in the clinical setting to evaluate lower extremity muscle groups. In the Thomas test position, hip flexors are considered within normal limits if the hip can be positioned in 0 degrees of flexion. Tensor fasciae latae and the ilio-tibial band can be tested in the Thomas position or Ober's position, with normal flexibility indicated by ability to position the hip in adduction or at least neutral abduction and adduction (Magee, 1987). Hip flexion of approximately 70 degrees with the knee extended is expected for hamstring flexibility. Dorsiflexion of 0 degrees with the knee extended and a

minimal increase of dorsiflexion with knee flexion are considered normal flexibility of the gastrocnemius and soleus muscles respectively (Donatelli & Walker, 1989). It must be stressed that association between muscle tightness and PFPS has been inferred from the results of treatment versus prospective population studies.

Variations in normal gait parameters are well known and thus the usefulness of gait evaluation, especially in measuring change in PFPS patient has been questioned. Chesworth et al. (1989) reported that temporal components of gait were not sensitive methods for measuring change in a small group of PFPS patients. The severity of pain was not considered in this study, and the authors suggest that with greater pain and a larger population gait changes may be evident. General clinical observations during gait may be useful for clarifying other findings in the lower extremity, however, until more information is available, clinicians should be cautious in the interpretation of gait observations in PFPS patients.

Radiological evaluation of PFPS patients has been shown to be primarily useful in diagnosing pathology such as: fractures, developmental abnormalities of the patella and femur, tumours and osteoarthritis. Many techniques have been described for assessing the patellar position and the patellofemoral articulation. A summary of the radiologic views, with normal and abnormal values, are clearly presented by Reid (1992). Standard anteroposterior, lateral and intra-patella roentgenographs at 30 and 90 degrees of knee flexion have been recommended by Goldberg (1991). Minkoff and Fein (1989) presented a detailed overview of the role of radiology in the management of PFPS, and recommended that stress films, magnetic resonance imaging, computerized tomography, isotope scanning and arthrography may be indicated for more serious cases. The relevance of radiographic findings to the presentation of the typical PFPS patient continues to be investigated.

Shellock et al. (1991) described an advancement in kinematic magnetic resonance imaging that has been successfully used to evaluate patellofemoral disorders. By taking multiple images at a temporal resolution suitable for knee

movement, the contribution of active muscle contraction to patellofemoral tracking and alignment was assessed. The authors concurred with other reports (Shellock, 1988; Kujala, Osterman, Kormano & Schlenzka, 1989; Schutzer, Ramsby & Fulkerson, 1986) that pathomechanics of the patellofemoral joint are prevalent at early stages of knee flexion. This brings into question the sensitivity of many of the techniques that position the knee in more flexion (ie. Hughston and Merchant tangential views).

Treatment

Conservative treatment for PFPS is advocated by many authors (Reid, 1991; McConnell, 1986; Andrews & Thornberry, 1986; Beckman, Craig & Lehman, 1989;Hughston, Walsh & Puddu, 1984), even in severe cases. Exercise, education, rest, bracing, wrapping, electrical modalities and anti-inflammatory medications have been components of many rehabilitation programs. The goals of conservative treatment include: reduction of the inflammatory process, increasing muscular control, reducing tightness of soft tissue and muscle structures, correction of biomechanical factors and modification of activity.

Most conservative treatments have been directed toward strengthening programs to increase the patellar stabilizing effects of the quadriceps muscle group, specifically the vastus medialis obliquus (VMO) muscle (Voight & Wieder, 1991). Methods for achieving this stabilizing effect while limiting excessive joint reaction force have consisted of isometric exercise with the knee in extension or slightly flexed (O'Neill, Micheli, Warner, 1992; Bentley & Dowd, 1984; Brunet & Stewart, 1989;Cerullo, Puddu, Conteduca, Ferretti, & Mariani, 1988; Kettelkamp, 1981; Kramer, 1986); multiple angle isometric knee extensor strengthening (Kramer, 1986); and terminal knee extension exercises (Brunet, 1989; Grana 1985). Eccentric quadriceps strengthening programs (Brunet, 1989; Bennett & Stauber, 1986; McConnell, 1986) and isokinetic exercises at high speeds (Steadman, 1979; Timm, 1988) have also been recommended. Many authors have incorporated hip adduction exercises into strengthening programs to facilitate VMO contractions (Antichi & Brewster, 1986; Carson, 1985; Hanten & Schulthies, 1990).

Flexibility programs have been considered essential components to PFPS rehabilitation by many authors. Hamstring, iliotibial band and gastrocnemius-soleus stretching exercises are most commonly advocated (Beckman, et al, 1989; Malek & Mangine, 1981; McConnell, 1986). Depending on the involvement of other muscle groups in the lower extremity and pelvic region more specific stretching and strengthening exercises may also be prescribed (McConnell, 1986). Endurance training has been recommended by several authors, however, caution must be exercised to ensure that symptoms are controlled. Bicycling and swimming are the most commonly selected activities because of their relatively low loading on the lower extremities (Ericson & Nisell, 1979; Malek & Mangine, 1981; Malone et al., 1980). Stair climbers have been suggested as advanced endurance activities (McConnell, 1986; Beckman et al., 1989).

Functionally oriented activities, emphasizing motor control of the quadriceps, especially VMO, have been more recently advocated (McConnell, 1986). Biofeedback has been used to facilitate muscle control and has been shown to change vastus medialis and vastus lateralis muscle activity (LeVeau & Rogers, 1980). Proprioception training has been included in some programs in order to improve kinesthetic sense in the knee and to introduce elements of functional activities (Malone, Blackburn & Wallace, 1980).

In reviewing the literature on PFPS management, Shelton and Thigpen (1991) reported the following trends in current treatment philosophies: comprehensive care based on thorough evaluation; dynamic patellar stability using exercise, biofeedback, education and functional activities; realignment of the patella using taping, mobilization techniques and bracing; use of medications, ice and other modalities for control of pain and inflammation; flexibility and strength programs aimed at reducing biomechanical faults in the lower extremity; and detailed patient education programs stressing long-term management and functional progression.

Arthroscopy may be used for diagnostic purposes and carrying out specific procedures such as retinacular releases. Varying success rates have been

reported following surgical intervention, thus surgery is usually delayed or avoided if possible (Gruber; 1979; Kettelkamp, 1981). Reid (1992) outlines surgical approaches for patellofemoral pain relative to five main categories of diagnosis: major tracking problems and dislocations; lateral facet hyperpressure; minor tracking problems and Grade I and II chondromalacia; symptomatic plica and Grade III and IV chondromalacia and arthrosis. Lateral retinacular and plical releases, patellar debridement, extensor mechanism realignment, facetectomy and patellectomy are some of the surgical techniques used in management of PFPS (Reid, 1991; Shelton & Thigpen, 1991; Andrews & Thornberry, 1986; Micheli & Stanitski, 1981; Paulos, et al., 1980).

Measuring Clinical Change

Clinical studies evaluating the effects of treatment intervention on PFPS patients have commonly reported changes in certain components of impairment and disability (O'Neill et al., 1992; Doucette & Goble, 1992; McConnell, 1986). Evaluation of various aspects of function, pain and specific objective clinical findings are commonly analyzed. Unfortunately many of these measurements have not been validated.

Although Noyes' (1990) Cincinnatti rating system was designed for ligamentous injuries of the knee, components of the scale have been recommended for the study of PFPS patients. Eight clinical parameters related to the patellofemoral joint have been included, and each factor has been assigned a rating. The degree of patellofemoral joint crepitus, level of pain with patellofemoral joint compression and soft tissue palpation, and degree of soft tissue swelling are four of the parameters evaluated. As well, the degree of lateral and medial patellar subluxation and measures of Q-angle at two different degrees of knee flexion are included in the scale (Appendix B-1).

An evaluative scale designed specifically for evaluating PFPS patients was developed by Reid (1992). This tool assessed and rated various aspects of the patient's history related to pain with activity, functional restrictions and the use of orthoses. Components of the physical examination were also included in Reid's

scale (Appendix B-2). Schwarz (1988) developed a patellofemoral rating scale for post-operative evaluation of osteochondritis dissecans patients. The factors included in this scale are similar to the one proposed by Reid (1992), with ratings arbitrarily assigned to each factor (Appendix B-3). Saltzman and colleagues. (1990) evaluated patients who had been treated for patellar fractures by partial patellectomy. The methods for evaluating the results of treatment included: a questionnaire based on Noyes' Cincinnatti system, radiographic assessment, and certain objective clinical findings. A cumulative score of 100 produced a patellofemoral index, which was analysed over a treatment period.

Validity has not been reported for the preceding tools. Also, in regards to the validity of objective clinical findings, studies have shown that the factors commonly evaluated, such as alignment, are not correlated with the severity of PFPS (Fairbank, et al., 1984; Wilson, 1985; Reikeras, 1992). Fairbank et al. (1984) reported that activity level was the only significant indicator for PFPS in adolescents and young adults. It must be noted however, that debate continues to surround the definitions of normal and abnormal values for many of the objective clinical tests.

Validation of outcome measures, including visual analogue scales for pain, a functional index questionnaire (FIQ) (Appendix C), and temporal gait components and EMG activity of quadriceps muscle during stair climbing, were investigated by Chesworth et al (1989). This study determined that visual analogue scales and the FIQ were useful tools for measuring change in clinical status, however EMG measurement and gait analysis were not sensitive methods for determining clinical change. Functional activities commonly affected in PFPS were represented in the FIQ. Walking, sitting for prolonged periods, squatting, kneeling, stairs and running were graded on an ordinal scale. Although changes in these functions have been frequently reported (Noyes, 1990, Reid, 1992, Saltzman et al.,1990), the weighting of each parameter in the various scales has been arbitrary. Validity of these tools on populations with different activity levels must also be evaluated, as activity level has been shown to greatly influence lower

extremity impairment (Noyes, 1990).

OBJECTIVES AND STAGES OF THE STUDY

The objectives of this study were:

1) to investigate the psychometric properties of existing PFPS clinical evaluation methods;

2) to determine the essential components of an evaluation tool by reviewing the literature;

3) to analyse the content validity of the components of PFPS evaluation;

4) to develop a PFPS evaluation tool that could be easily implemented into a clinical environment;

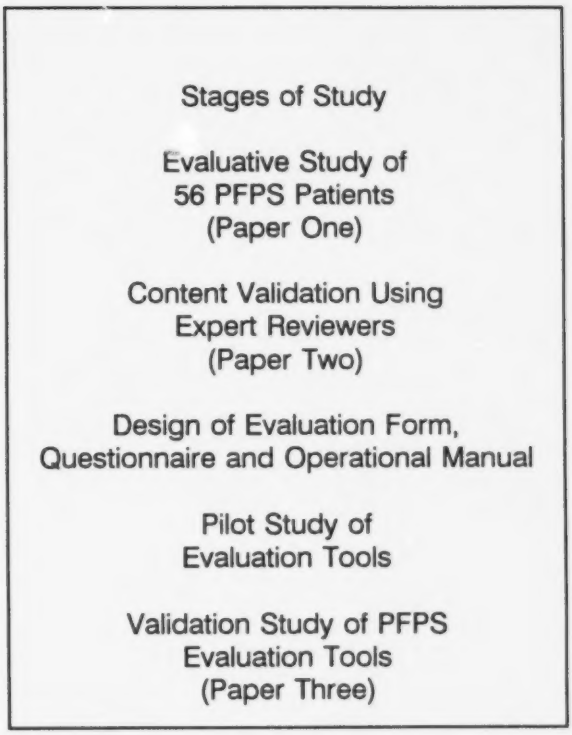
5) to analyse measurement properties of the different components of the evaluation tool including: reliability, criterion validity, construct validity, sensitivity, specificity and effect size.

The development of the PFPS clinical evaluation methods consisted of four stages. Stage one investigated the effectiveness of several outcome measures in detecting clinical change in a population of 56 PFPS patients. The patients participated in a randomized controlled clinical trial evaluating the efficacy of three treatment programs, with measures taken prior to treatment and after one month of treatment. The methodology and results are presented in Paper One " Analysis of Outcome Measures used in the Study of PFPS". Stage two of development consisted of a process of content validation. As well as reviewing the literature on evaluation methods in PFPS, a questionnaire (Appendix D) was designed to survey expert reviewers' opinions on content of an evaluation tool. The methodology and results of this process are detailed in Paper Two "Development of a Clinical Tool and Patient Questionnaire for the Evaluation of PFPS Patients".

The third stage of the study consisted of development of: a PFPS clinical evaluation form (Appendix E), a patient questionnaire (Appendix F), an operational manual to standardize evaluation procedures and recording (Appendix G), and an instructional videotape including the contents of the operational manual. These

components were reviewed and piloted using four physical therapists with limited experience in treating PFPS. It was felt that if clinicians with limited exposure to PFPS patients could easily interpret and use the evaluation tools, that this would increase utility of the tool in a variety of clinical settings. Three subjects were evaluated during one session, in random order by each of the physical therapists. The three subjects included two females diagnosed with PFPS and one male with no reported knee problems. Physical therapists were interviewed following the assessments and feedback regarding the clarity of definitions and evaluation format were provided. In addition, each subject completed a questionnaire and provided feedback on clarity of wording and format. Based on these results, minor revisions were made to the evaluation form, operational definitions and patient questionnaire.

The final stage of the study investigated the validity of the PFPS evaluation tool. Physical therapists from seven clinical facilities in western Canada volunteered to participate in the study. Of these facilities, four facilities were able to provide patients over the six month study period. PFPS patients and non-PFPS subjects were recruited and tested on two occasions. Methodology and results of this study are reported in Paper Three "Validation of a Clinical Evaluation Tool for PFPS".


Stages of Study
Evaluative Study of
56 PFPS Patients
(Paper One)
Content Validation Using
Expert Reviewers
(Paper Two)
Design of Evaluation Form,
Questionnaire and Operational Manual
Pilot Study of
Evaluation Tools
Validation Study of PFPS
Evaluation Tools
(Paper Three)

BIBLIOGRAPHY

Abernathy PJ, Townsend PR, Rose RM, Radin EL: Is chondromalacia a separate clinical entity? J Bone Joint Surg [Br] 1978; 60: 205-210.

Aday LA: Designing and Conducting Health Surveys. 1st ed. San Francisco: Jossey-Bass Inc., 1989.

Ainsworth BE, Jacobs DR, Leon AS: Validity and reliability of self-reported physical activity status: The Lipid Research Clinics questionnaire. Med Sci Sports Ex 1993; 125(No.1): 92-98.

American Psychological Association: Standards for Education and Psychological Tests. Washington,D.C.: American Psychological Assoc., 1974.

Andersson G: Hip assessment: A comparison of nine different methods. J Bone Joint Surg [Br] 1972; 54-B(4): 621-625.

Antichi TJ, Brewster CE: Modification of quadriceps femoris muscle exercises during knee rehabilitation. Phys Ther 1986; 66: 1246-1251.

Andriacchi TP, Andersson GB, Fermier RW, Stern D, Galante JO: A study of lower-limb mechanics during stair climbing. JBJS[Am} 1980; 62-A(5): 749-757.

Arno S: The A angle: A quantitative measurement of patella alignment and realignment. JOSPT 1990; 12(6): 237-242.

Baird JC, Noma E: Fundamentals of Scaling and Psychophysics. New York: Wiley, 1978.

Barber SD, Noyes FR, Mangine RE, McCloskey JW, Hartman W: Quantitative assessment of functional limitations in normal and anterior cruciate ligament-deficient knees. The Noyes Knee Rating System. An Internal Publication of Cincinnati Sportsmedicine Research and Education Foundation, Cincinnati, Ohio, 1990.

Beckman M, Craig R, Lehman RC: Rehabilitation of patellofemoral dysfunction in the athlete. Clin Sports Med 1989; 8: 841-860.

Begner M, Bobbitt RA, Pollard WE, Martin DP, Gilson BS: The sickness

impact profile: validation of a health status measure. MedCare 1976; 14: 57-67.

Bennett JG, Stauber WT: Evaluation and treatment of anterior knee pain using eccentric exercise. Med Sci Sports Ex 1986; 18(5): 526-530.

Bentley G, Dowd G: Current concepts of etiology and treatment of chondromalacia patellae. Clin Orthop 1984; 189: 209-228.

Berg K, Wood-Dauphinee, Williams JI, Gayton D: Measuring balance in the elderly: Preliminary development of an instrument. PhysiotherCan 1989; 41: 304-311.

Bose K, Kanagasuntheram R, Osman MBH: Vastus medialis oblique: an anatomic and physiologic study. Orthopedics 1980; 3: 880-883.

Bostram C, Harms-Ringdahl K, Hordemar R: Clinical reliability of shoulder function assessment in patients with rheumatoid arthritis. ScanJRheumatology 1991; 20: 36-48.

Boureau F, Luu M, Doubrere JF: Comparative study of the validity of four French McGill Pain Questionnaire (MPQ) versions. Pain 1992; 50: 59-65.

Bourne MH, Hazel WA, Scott SA, Sim FH: Anterior knee pain. MayoClinProc 1988; 63: 482-491.

Brunet ME, Stewart GW: Patellofemoral rehabilitation. Clin Sports Med 1989; 8: 319-329.

Budiman-Mak E, Conrad KJ, Roach KE: The foot function index: A measure of foot pain and disability. JClinEpidemiol 1991; 44: 561-570.

Campbell DT, Fiske DW: Convergent and discriminant validation by the multitrait multimethod matrix. PsycholBull 1959; 56: 81-105.

Carson WG: Diagnosis of extensors mechanism disorders. Clin Sports Med 1985; 4: 235-246.

Cerullo G, Puddu G, Conteduca F, Ferretti A, Mariani PP: Evaluation of the results of extensor mechanism reconstruction. Am J Sports Med 1988; 16: 93-96.

Chambers LW, MacDonald LA, Tugwell P, Buchanan WW, Kraag EG: The McMaster health index questionnaire as a measure of quality of life for

patients with rheumatoid disease. JRheumatol 1982; 9: 780-783.

Chesworth BM, Culham EG, Tata GE, Peat M: Validation of outcome measures in patients with patellofemoral syndrome. JOSPT 1989; 11: 302-308.

Clark WC, Yang JC: Application of Sensory Decision Theory to Problems in Laboratory and Clinical Pain. In: Melzack R, ed. Pain Measurement and Assessment. New York: Raven Press, 1983:15-25.

Cohen J: Statistical Power Analysis for the Social Sciences. 2nd ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1988.

Craig KD, Prkachin KM: Nonverbal Measures of Pain. In: Melzack R, ed. Pain Measurement and Assessment. New York: Raven Press, 1983:173-179.

Cronbach LJ: Test Validation. In: Thorndike RL, ed. Educational Measurement. 2nd ed. Washington,D.C.: American Council on Education, 1971.

Cyriax J: Textbook of Orthopedic Medicine. Vol1, 6ed. New York: MacMillan Publishing Co., 1975.

DeHaven DE, Dolon WA, Mayer PJ: Chondromalacia patellae in athletes. Clinical presentation and conservative management. Am J Sports Med 1979; 7: 5-11.

Delitto A: Subjective measures and clinical decision making. Phys Ther 1989; 69: 585-589.

Devereaux MD, Lachmann SM: Patello-femoral arthralgia in athletes attending a sports injury clinic. Br J Sports Med 1984; 18(1): 18-21.

Deyo RA: Measuring the functional status of patients with low back pain. Arch Phys Med Rehabil 1988; 69: 1044-1053.

Deyo RA, Centor RM: Assessing the responsiveness of functional scales to clinical change: An analogy to diagnostic test performance. JChronDis 1986; 39: 897-906.

Deyo RA, Inui TS: Towards clinical applications of health status measures: Sensitivity of scales to clinically important changes. Health Serv Res 1984; 19: 275-289.

Deyo RA, Patrick DL: Barriers to the use of health status measures in clinical investigation, patient care, and policy research. MedCare 1989; 27(3 Suppl. 1): S254-S268.

DiVeta JA, Vogelback WD: The clinical efficacy of the A-angle in measuring patellar alignment. JOSPT 1992; 16(3): 136-139.

Dixon JS, Bird HA: Reproducibility along a 10cm vertical visual analogue scale. AnnRheumDis 1981; 40: 87-89.

Donatelli R, Walker R: Lower quarter evaluation: Structural relationships and interdependence. In: Donatelli R, Wooden J eds. Orthopaedic Physical Therapy New York: Churchill Livingstone, 1989: 277-304.

Doucette SA, Goble EM: The effect of exercise on patellar tracking in lateral patellar compression syndrome. Am J Sports Med 1992; 20(4): 434-440.

Downie WW, Leatham PA, Rhind VM, Wright V, Brancho JA, Anderson JA: Studies with pain rating scales. AnnRheumDis 1978; 37: 378-381.

Doxey GE: Girth measurements and muscle bulk. JOSPT 1987; 9(5): 177-183.

Dugdale TW, Barnett PR: Historical background: Patellofemoral pain in young people. OrthopClinNAm 1986; 17: 211-219.

Eisele SA: A precise approach to anterior knee pain: More accurate diagnoses and specific treatment programs. Physician Sportsmed 1991; 19(6): 126-130.

Ericson MO, Nisell R: Patellofemoral joint forces during ergometric cycling. Phys Ther 1987; 67: 1365-1369.

Fairbank JCT, Pynsent PB, van Poortvliet JA, Phillips H: Mechanical factors in the incidence of knee pain in adolescents and young adults. J Bone Joint Surg [Br] 1984; 66B: 685-693.

Feinstein AR: Clinimetrics. New Haven CT: Yale Univ Press, 1987.

Ferketich SL, Figueredo AJ, Knapp TR: Focus on psychometrics The Multitrait-Multimethod Approach to construct validity. ResNurs& Health 1991; 14: 315-320.

Ficat RP, Hungerford DS: Disorders of the Patello-Femoral Joint. Baltimore:

Williams & Wilkins, 1977.

Figueredo AJ, Ferketich SL, Knapp TR: Focus on psychometrics More on MTMM: The role of confirmatory factor analysis. ResNurs& Health 1991; 14: 387-391.

Flandry F, Hunt JP, Terry GC, Hughston JC: Analysis of subjective knee complaints using visual analogue scales. Am J Sports Med 1991; 19(2): 112-118.

Fletcher RH, Fletcher SW, Wagner EH: Chapter 3 Diagnosis In: Clinical Epidemiology. 2nd Ed. Baltimore: Williams & Wilkins, 1988: 42-75.

Fordyce WE: The Validity of Pain Behavior Measurement. In: Melzack R, ed. Pain Measurement and Assessment. New York: Raven Press, 1983:145-153.

Fries JF, Spitz PW, Young DY: The dimensions of health outcomes: The health assessment questionnaire, disability and pain scales. JRheumatol 1982; 9: 789-793.

Fulkerson JP, Hungerford DS: Disorders of the Patellofemoral Joint. 2nd ed. Baltimore: Williams and Wilkins, 1990.

Gaito J: Measurement scales and statistics: Resurgence of an old misconception. PsycholBull 1982; 87: 504-507.

Gaughan AM, Gracely RH, Friedman R: Pain perception following regular practice of meditation, progressive muscle relaxation and sitting. Pain 1990(Suppl.5): S31.

Goldberg B: Chronic anterior knee pain in the adolescent. PediatrAnn 1991; 20: 186-7,190-3.

Goodfellow J, Hungerford DS, Zindel M: Patellofemoral joint mechanics and pathology. J Bone Joint Surg [Br] 1976; 58(3): 287-290.

Goodwin LD, Goodwin WL: Focus on psychometrics Estimating construct validity. ResNurs& Health 1991; 14: 235-243.

Gouveia WA, Bungay KM, Massaro FJ, Ware JE: Pardigm for the management of patient outcomes. AJHP 1991; 48: 1912-1916.

Gracely RH: Evaluation of multi-dimensional pain scales. Pain 1992; 48: 297-300.

Grandjean E: Fitting the task to the man. An ergonomic approach. London: Taylor & Francis Ltd., 1982. (pp.65-71)

Granger CV, McNamara MA: Functional Assessment Utilization: The Long-Range Evaluation System (LRES). In: Granger C, Gresham G, eds. Functional Assessment in Rehabilitation Medicine. Baltimore: Williams & Wilkins, 1984:22-121.

Gruber MA: The conservative treatment of chondromalacia patella. OrthopClinNAm 1979; 10: 105-115.

Guccione AA: Physical therapy diagnosis and the relationship between impairments and function. Phys Ther 1991; 71: 499-503.

Guyatt G, Walter S, Norman G: Measuring change over time: assessing the usefulness of evaluative instruments. JChronDis 1987; 40: 171-178.

Guyatt GH, Deyo RA, Charlson M, Levine MN, Mitchell A: Responsiveness and validity in health status measurement: a clarification. JClinEpidemiol 1989; 42: 403-408.

Hanten WP, Schulthies SS: Exercise Effect on Electromyographic Activity of the Vastus Medialis Oblique and Vastus Lateralis Muscles. Phys Ther 1990; 70: 561-565.

Heft MW, Gracely RH, Dubner R: Nitrous oxide analgesia: A psychophysical evaluation using verbal descriptor scaling. JDentRes 1984; 63: 129-132.

Hefzy MS, Jackson WT, Saddemi SR, Hsieh Y-F: Effects of tibial rotations on patellar tracking and patello-femoral contact areas. JBiomedEng 1992; 14: 329-343.

Helmstader GC: Principles of Psychological Measurement. New York: Appleton-Century Crofts, 1964.

Henry JH, Goletz TH, Williamson B: Lateral retinacular release in patellofemoral subluxation. Am J Sports Med 1986; 14(2): 121-129.

Hoke B, Howell D, Stack M: The relationship between isokinetic testing and dynamic patellofemoral compression. JOSPT 1983; 4: 150.

Holden, Fekken, Jackson: Structured personality test item characteristics and validity. Journal of Research in Personality 1985; 19: 386-394.

Holroyd KA, Holm JE, Keefe FJ, et al: A multi-center evaluation of the McGill pain questionnaire: results from more than 1700 chronic pain patients. Pain 1992; 48: 301-311.

Hughston JC, Walsh WM, Puddu G: Patellar Subluxation and Dislocation. Philadelphia: WB Saunders, 1984.

Huskisson EC: Measurement of pain. Lancet 1974; 2: 1127-1131.

Huskisson EC: Measurement of Pain. JRheumatol 1982; 9: 768-769.

Huskisson EC: Visual Analogue Scales. In: Melzack R, ed. Pain Measurement and Assessment. New York: Raven Press, 1983:33-37.

Huskisson EC, Jones J, Scott PJ: Application of visual-analogue scales to the measurement of functional capacity. Rheumatol Rehabil 1976; 15: 185-187.

Insall J: Chondromalacia patella: Patellar malalignment syndrome. OrthopClinNAm 1979; 10: 117-127.

Insall J, Salvalti E: patella position in the normal knee joint. Radiol 1971; 101: 101-104.

Jette AM: Measuring subjective clinical outcomes. Phys Ther 1989; 69: 580-584.

Johanson NA, Charlson ME, Szatrowski TP, Ranawat CS: A self-administered hip-rating questionnaire for the assessment of outcome after total hip replacement. J Bone Joint Surg [Am] 1992; 74-A(4): 587-597.

Kaplan RM, Bush JW, Berry CC: Health status: types of validity and the index of well-being. Health Serv Res 1976; 11: 478-507.

Kazis LE, Anderson JJ, Meenan RF: Effect sizes for interpreting changes in health status. MedCare 1989; 3(Suppl. 1): S178-S189.

Kerlinger FN: Foundations of Behavioral Research. 2nd ed. New York: Holt,Rinehart and Winston, 1973.

Kettelkamp DB: Current concepts review: Management of patellar malalignment. J Bone Joint Surg [Am] 1981; 63A: 1344-1348.

Kibler WB: Strength and flexibility findings in anterior knee pain syndrome

in athletes. Proceedings of the International Society of the Knee. Sydney, Australia, 1987.

Kirshner B, Guyatt G: A methodological framework for assessing health indices. JChronDis 1985; 38: 27-36.

Kolowich PA, Paulos LE, Rosenberg TD, Farnsworth S.: Lateral release of the patella: Indications and contraindications. Am J Sports Med 1990; 18: 359-365.

Kramer PG: Patellar malalignment syndrome: Rationale to reduce excessive lateral pressure. JOSPT 1986; 8: 301-309.

Kujala UM, Osterman K, Kormano M, Komu M, Schlenzka D: Patellar motion analyzed by magnetic resonance imaging. Acta Orthop Scand 1989; 60: 13-16.

Larson JS: The Measurement of Health. 1st ed. New York: Greenwood Press, 1991.

Laurin CA, Levesque HP, Dussault R, LaBelle H, Peides JP: The abnormal lateral patellofemoral angle. J Bone Joint Surg [Am] 1978; 60-A: 55-60.

Leavitt F: Detecting psychological disturbances using verbal pain measurement: The Back Pain Classification Scale. In: Melzack R, ed. Pain Measurement and Assessment. New York: Raven Press, 1983:79-84.

Leslie IJ, Bentley G: Arthroscopy in diagnosis of chondromalacia patellae. AnnRheumDis 1978; 37: 540-547.

LeVeau BF, Roger C: Selective training of the vastus medialis muscle using EMG biofeedback. Phys Ther 1980; 60: 1410-1415.

Liang MH, Fossel AH, Larson MG: Comparisons of five health status instruments for orthopedic evaluation. MedCare 1990; 28: 632-642.

MacIntyre D, Wessel J: Knee muscle torques in patellofemoral pain syndrome. PhysiotherCan 1988; 40: 20-24.

MacIntyre DL, Robertson DGE: Quadriceps Muscle Activity in Women Runners With and Without Patellofemoral Pain Syndrome. Arch Phys Med Rehabil 1992; 73: 10-14.

Magee DJ: Orthopedic Physical Assessment. 1st ed. Philadelphia:

W.B.Saunders Co., 1987.

Malek MM, Mangine RE: Patellofemoral pain syndromes: a comprehensive and conservative approach. JOSPT 1981; 2: 108-116.

Malgady RG, Krebs DE: Understanding correlation coefficient and regression. Phys Ther 1986; 66: 110-119.

Malone T, Blackburn TA, Wallace LA: Knee rehabilitation. Phys Ther 1980; 60: 1602-1610.

Manne SL, Jacobsen PB, Redd WH: Assessment of acute pediatric pain: do child self-report, parent ratings, and nurse rating measure the same phenomenon? Pain 1992; 48: 45-52.

McConnell J: The management of chondromalacia patellae: A long term solution. AustralJPhysiother 1986; 32: 215-223.

McDowell I, Newell C: Measuring Health A Guide to Rating Scales and Questionnaires. New York: Oxford University Press, 1987.

McKenzie CR, Charlson ME: Standards for the use of ordinal scales in clinical trials. BMedJ 1986; 292: 40-43.

McMullen W, Roncarati A, Koval P: Static and isokinetic treatments of chondromalacia patella: A comparative investigation. JOSPT 1990; 12: 256-266.

Meenan RF, Gertman PM, Mason JH: Measuring health status in arthritis. Arthritis Rheum 1980; 23: 146-152.

Melzack R, ed: Pain Measurement and Assessment. New York: Raven Press, 1980.

Melzack R, ed: Pain Measurement and Assessment. New York: Raven Press, 1983.

Melzack R: Letter to the Editor. Pain 1985; 23: 201-203.

Melzack R, Wall PD: Pain mechanisms: A new theory. Science 1965; 150: 971-979.

Merchant AC, Mercer RL, Jacobsen RH, Cool CR: Roentgenographic analysis of patellofemoral congruence. J Bone Joint Surg [Am] 1974; 56-A:

1391-1396.

Metcalf R: An arthroscopic method for lateral release of the subluxating or dislocating patella. Clin Orthop 1982; 167: 9.

Micheli LJ, Stanitski CL: Lateral patellar retinacular release. Am J Sports Med 1981; 9: 330-335.

Milgrom C, Kerem E, Finestone A, Eldad A, Shlamkovitch N: Patellofemoral pain caused by overactivity. J Bone Joint Surg [Am] 1991; 73-A(7): 1041-1043.

Minkoff J, Fein L: The role of radiography in the evaluation and treatment of common anarthrotic disorders of the patellofemoral joint. Clin Sports Med 1989; 8: 203-260.

Minns RJ, Birnie AJ, Abernathy PJ: A stress analysis of the patella and how it relates to the patellar articular cartilage lesions. JBiomech 1979; 12: 699-711.

Moller BN, Jurik AG, Tidemand-Dal C, Krebs B, Aaris K: The quadriceps function in patellofemoral disorders. Arch Orthop Trauma Surg 1987; 106: 195-198.

Nissel R, Ericson M: Patellar forces during isokinetic knee extension. Clin Biomech 1992; 7: 104-108.

Nordin M, Frankel BH: Basic Biomechanics of the Musculoskeletal System. 2nd ed. Philadelphia: Lea & Febiger, 1989.

Norman GR: Issues in the use of change scores in randomized trials. JClinEpidemiol 1989; 42: 1097-1105.

Noyes FR: The Noyes Knee Rating System An assessment of subjective, objective, ligamentous, and functional parameters. Internal Publication of Cincinnati Sportsmedicine Research and Education Foundation. Cincinnati, Ohio (1990).

Noyes FR, Barber SD, Mangine RE: Abnormal lower limb symmetry determined by function hop test after anterior cruciate ligament rupture. Am J Sports Med 1991; 19(5): 516-518.

Noyes FR, Barber SD, Mooar LA: A rationale for assessing sports activity levels and limitations in knee disorders. ClinOrthopRelRes 1989; 246:

238-249.

Noyes FR, Mooar LA, Barber SD: The assessment of work-related activities and limitations in knee disorders. Am J Sports Med 1991; 19: 178-188.

Noyes FR, Stabler CL: A system for grading articular cartilage lesions at arthroscopy. Am J Sports Med 1989; 17(4): 505-513.

Nunally JC: Psychometric Theory. New York: McGraw-Hill Book Co., 1967.

Nunally JC: Introduction to Psychological Measures. New York: McGraw-Hill Book Co., 1970.

O'Neill DB, Micheli LJ, Warner JP: Patellofemoral stress. Am J Sports Med 1992; 20(2): 151-156.

Outerbridge RF: The problem of chondromalacia patellae. Clin Orthop 1975; 110: 177.

Paulos L, Rusche K, Johnson C, Noyes FR: Patellar malalignment: A treatment rationale. Phys Ther 1980; 60: 1624-1632.

Payton OD: Research: The Validation of Clinical Practice. 2nd ed. Philadelphia: F.A.Davis Co., 1988.

Reid DC: Sports Injury Assessment and Rehabilitation. New York: Churchill Linvingstone, 1992.

Reikeras D: Patellofemoral characteristics in patients with increased femoral anteversion. SkelRadiol 1992; 21: 311-313.

Riddle DL, Finucane SD, Rothstein JM, Walker ML: Intrasession and intersession reliability of hand-held dynamometer measurements taken on brain-damaged patients. Phys Ther 1989; 69: 182-194.

Robert JM: The effect of taping on patellofemoral alignment - A radiological pilot study. In: Proceedings, Sixth Biennial Conference Manipulative Therapists Assocation of Australia, Melbourne 1989: 146-151.

Rothstein JM, ed: Clinics in Physical Therapy: Measurement in Physical Therapy. v. 7. New York: Churchill Livingstone, 1985.

Saltzman CL, Goulet JA, McClellan RT, Schneider LA, Matthews LS: Results of treatment of displaced patellar fractures by partial patellectomy. J Bone

Joint Surg [Am] 1990; 72-A(9): 1270-1284.

Sanders SH: Toward a practical instrument system for the automatic-measurement of up-time in chronic pain patients. Pain 1980; 1: 103-109.

Schutzer SF, Ramsby GR, Fulkerson JP: Computed tomographic classification of patellofemoral joint pain patients. OrthopClinNAm 1986; 17: 235-248.

Schwarz C, Blazina ME, Siston DJ, Hirsh LC: The results of operative treatment of osteochondritis dissecans of the patella. Am J Sports Med 1988; 16(5): 522-529.

Scott J, Huskisson EC: Accuracy of subjective measurements made with or without previous scores: An important source of error in serial measurement of subjective states. AnnRheumDis 1979; 38: 558-559.

Shellock FG, Foo TKF, Deutsch AL, Mink JH: Patello-femoral joint: Evaluation during active flexion with ultrafast spoiled GRASS MR imaging. Radiol 1991; 180: 581-585.

Shellock FG, Mink JH, Fox JM: Patellofemoral joint: kinematic MR imaging to assess tracking abnormalities. Radiol 1988; 168: 551-553.

Shelton GL, Thigpen LK: Rehabilitation of patellofemoral dysfunction: A review of literature. JOSPT 1991; 14(6,Dec): 243-249.

Shrout PE, Fleiss JL: Intraclass correlations: Uses in assessing rater reliability. PsycholBull 1979; 86: 420-428.

Smith GT, ed: Measuring Health: A Practical Approach. Chichester: John Wiley & Sons, 1988.

Smythe HA, Helewa A, Goldsmith CH: Selection and combination of outcome measures. JRheumatol 1982; 9: 770-773.

Souza DR, Gross MT: Comparison of vastus medialis obliquus: vastus lateralis muscle integrated electromyographic ratios between healthy subjects and patients with patellofemoral pain. Phys Ther 1991; 71: 310-316.

Steadman JR: Nonoperative measures for patellofemoral problems. Am J Sports Med 1979; 7: 374-375.

Straub T, Hunter RE: Acute anterior cruciate ligament repair. Clin Orthop 1988; 227: 238.

Stratford P: A Clinical Measurment Study Design Guide Draft 2 Prepared in cooperation with Research Division of Canadian Physiotherapy Association (1991).

Streiner DL, Norman GR: Health Measurement Scales. 1st ed. Oxford: Oxford Medical Publications, 1989.

Tegner Y, Lysholm J: Rating systems in the evaluation of knee ligament injuries. Clin Orthop 1985; 198: 43-49.

Timm KE: Postsurgical knee rehabilitation. Am J Sports Med 1988; 16: 463-468.

Tsirbas A, Paterson RS, Keene GCR: Fat pad impingement; A missed cause of patello-femoral pain? The Australian Journal of Science and Medicine in Sport 1990: 24-26.

Tugwell P, Sinclair A, Dok C, Anderson G, Buchanan WW: Preference for endpoint measures in clinical trials: results of structured workshops. JRheumatol 1982; 9: 798-801.

Tuley MR, Mulrow CD, McMahan CA: Estimating and testing an index of responsiveness and the relationship of the index to power. JClinEpidemiol 1991; 44: 417-421.

Voight ML, Wieder DL: Comparative reflex response times of vastus medialis obliquus and vastus lateralis in normal subjects and subjects with extensor mechanism dysfunction. Am J Sports Med 1991; 19(2): 131-137.

Waisbrod H, Treiman N: Intra-osseous venography in patellofemoral disorders: a preliminary report. J Bone Joint Surg [Br] 1980; 62: 454-456.

Wiberg G: Roentgenographic and anatomic studies on the femoropatellar joint with special reference to chondromalacia patellae. Acta Orthop Scand 1941; 12: 319-409.

Wilson DJ: Treatment of anterior knee pain by calf stretching and walking re-education. Physiotherapy 1989; 75: 127-128.

Wilson JM: Interosseus pressures of the patella in chondromalacia [Dissertation]. University of Alberta, 1985.

Wilson T: Anterior knee pain: A new technique for examination and treatment. Physiotherapy 1990; 76: 371-376.

Wittenborn JR: Reliability, validity and objectivity of symptom-rating scales. JNervMentDis 1972; 154: 79-87.

Wonca Classification Committee, ed: Frontiers of Primary Care: Functional Status Measurement in Primary Care. New York: Springer-Verlag, 1990.

World Health Organization: International classification of impairments, disabilities and handicaps. A manual of classification relating to the consequences of disease. Geneva: World Health Organization, 1980.

Wright JG, Feinstein AR: Improving the reliability of orthopaedic measurements. J Bone Joint Surg [Br] 1992; 74-B(2): 287-291.

Young A, Hughes I, Round P, Edwards PJR: Measurement of quadriceps muscle wasting by ultrasonography. Rheumatol Rehabil 1980; 17: 141-147.

CHAPTER TWO

PAPER ONE

ANALYSIS OF OUTCOME MEASURES USED IN THE STUDY OF PATELLOFEMORAL PAIN SYNDROME

INTRODUCTION

The tools or clinical methods used to evaluate change in physical status of patients have been progressively scrutinized over the past years (McDowell & Newell, 1987). Without appropriate tools the clinician is unable to effectively monitor treatment programs and patient response. Selection of the essential components that should comprise any measurement tool is influenced by the individual, the medical condition, and the environment. Theoretically, every person should be evaluated relative to his or her specific problems, with changes in symptoms and signs considered as relevant indicators on which to base treatment. The single case design is based on this philosophy, and support for this type of research has been growing. Population based research is also necessary, in order that basic assessment and treatment methods can be developed as a foundation on which to educate clinicians and determine the most effective treatment programs for patient populations. This type of research has a bearing on the delivery and funding of health care services.

Clinimetrics, as discussed by Delitto (1989), involves the quantification of clinical data and considers four features as essential to the development of appropriate measurements. The measure must demonstrate face validity, content validity, ease of usage and suitability for the clinical environment. Face validity is considered to be a low level of validity as it reflects only whether a measure appears to do what it is intended to do. Content validity concerns whether the measurements adequately reflect the variable that has been defined. In addition, the measure must be replicable and responsive to change. In evaluation of diagnostic tests, the relationship between the test and the presence of the disease or condition are often presented by describing the sensitivity (ie. proportion of true positive values detected by the test) and the specificity (ie. proportion of true negative values detected by the test). A sensitive test is used when there is an important penalty for missing a disease; and is used to rule out diseases (ie. establish that certain diseases are unlikely possibilities). Specific tests are especially important when false positive tests can be harmful to the patient

physically, emotionally or financially. Often they are used to "rule in " a diagnosis suggested by other data (Fletcher, Fletcher & Wagner, 1988).

In the development of functional status measurements, clinical endpoints are often combined to develop a single outcome score (Smythe, Helewa & Goldsmith, 1982). When a battery of tests are used together, it is referred to as a measurement index. An evaluative index is used to measure the magnitude of the change over time in an individual or group (Kirschner & Guyatt, 1985). Caution must be exercised in the development of indices, as information may be lost when combined results are analysed. Also, the combination of several clinical endpoints may not reflect functional status (Deyo & Patrick, 1989). Development of tools that include fundamental components of function, relevant to the entire population, is very complex, as activity levels and methods of performing functional activities vary greatly within the healthy population.

The management of patellofemoral pain syndrome (PFPS) is a challenge to the clinician, as the problem varies in intensity and duration and is usually activity specific. PFPS is a condition that is characterized by pain associated with activities that require movement at the patellofemoral joint during lower extremity activities. The condition is commonly reported in adolescent populations and young adults, with an insidious onset of vague aching pain over the peripatellar region. Often the pain is bilateral and aggravated by activity, especially stairs, squatting and prolonged sitting (Reid, 1992; Goldberg, 1991). Medical terminology for this condition varies, with the following descriptions commonly used: patellagia, anterior knee pain, peripatellar syndrome, malalignment syndrome, patellar arthralgia and chondromalacia patella. The literature on diagnosis and management of patellofemoral conditions is extensive. The prevalence of this condition in the adolescent population has been reported as high as 30% (Fairbank, Pynsent, vanPoortvliet & Phillips, 1984), while in the athletic population 25% of participants may complain of this condition (McConnell, 1986).

The exact etiology of the condition has not been determined and it is quite apparent that more than one cause is responsible for this condition. Eisele (1991)

reported that the most common etiologies included: patellar tracking disorders and instability; tendinitis; bursitis; synovitis; and traction apophysitis (Osgood Schlatter's Diseases). Most of the recent literature has categorized PFPS according to the presence of articular cartilage damage. In addition to the status of the articular cartilage, other factors that are frequently considered in classification are: proximal and distal biomechanical relationships, such as excessive sub-talar pronation and knee hyperextension; trauma, and local vasculature (Reid, 1992; Wilson, 1990; Bourne, Hazel, Scott & Sim, 1988; Paulos, Rusche, Johnson & Noyes, 1980; Ficat & Hungerford, 1977). Abnormal tracking appears to be one of the underlying causes of PFPS frequently discussed (Reid, 1992; McConnell, 1986; Insall, 1979). Changes in ligamentous and soft tissue stabilizers or the osseous structures of the patella and femur, along with muscle imbalance of the different components of the quadriceps, may result in the abnormal distribution of forces at the patellofemoral joint. As articular cartilage does not have nerve endings, the source of pain is unclear. Pain theories presently suggest that the pain sources may include: subchondral bone, synovium, capsule and venous engorgement (Minns, Birnie & Abernathy, 1979; Abernathy, Townsend, Rose & Radin, 1978; Waisbrod & Treiman, 1980).

Clinical studies evaluating the effects of treatment on PFPS patients have reported certain impairments, such as decreased muscle strength, and to a limited degree some disabilities, such as inability to climb stairs (O'Neill, Micheli & Warner, 1992; Doucette & Goble, 1992; McConnell, 1986). Various aspects of function, pain and objective clinical findings are commonly analysed (Chesworth, Culham, Tata & Peat, 1989). Reid (1992) presented an index for evaluating PFPS patients, which rated parts of the physical examination and aspects of the patient's history related to pain, activity, functional restrictions, and use of orthoses. Schwarz (1988) developed a patellofemoral rating scale for post-operative evaluation of patients with osteochondritis dissecans occurring at the knee. Based on Noyes (1990) Cincinnatti Knee Rating system, a patellofemoral index was developed (Saltzman, Goulet, McCelellan et al., 1990) to evaluate patients over a treatment

period. None of these tools have been evaluated for properties of validity.

Clinical studies have shown that factors commonly associated with PFPS patients, such as alignment, are poorly correlated with the severity of the condition (Reikeras, 1992; Wilson, 1985; Fairbank et al., 1984). Fairbank et al. (1984) found activity level to be the only significant indicator correlated with PFPS in adolescents and young adults. An evaluative study investigated the validity of: visual analogue scales for assessing pain, a functional index questionnaire, temporal gait components and electromyographic (EMG) activity of the quadriceps muscle group during stair climbing, in PFPS patients (Chesworth et al., 1989). Chesworth et al. (1989) recommended that visual analogue measurements of pain and the functional index questionnaire were effective in measuring changes in pain and function in PFPS patients, however gait and EMG variables did not change despite improvement in patients' conditions.

The goals of the present study were to: 1) determine the test-retest reliability of three evaluation tools; 2) analyse internal consistency of a questionnaire and a function scale; and 3) estimate the size and direction of change for each of the measures. The tools or measures evaluated in this study included: a functional index questionnaire (FIQ) (Chesworth et al., 1989), visual analogue scales (VAS) for pain, a patellofemoral function scale (PFS) (Reid, 1992), a step test, and subjects' judgement of functional limitations and self-reporting of change in condition following treatment.

METHODOLOGY

Fifty-six subjects participating in a one year prospective evaluative study on PFPS were included. The purpose of the evaluative study was to determine the efficacy of three different treatment approaches for PFPS patients. Subjects were referred by physicians to the study if they met two of the following criteria: a positive Clarke's sign, tenderness on palpation of the medial or lateral facets of the patella, or patellofemoral pain with isometric contraction of the quadriceps muscle group against minimal manual resistance with the knee positioned at any degree of normal movement. Exclusion criteria included: history of dislocation or

subluxation; current or previously diagnosed ligamentous, meniscal, fat pad, tendon, or bursae involvement; gross knee effusion; referred pain from the spine or hip; radiological changes associated with arthritis, osteochondral or chondral fractures of the knee; history of steroid injections to the knee; history of knee surgery or pending knee surgery; and any evidence or history of an upper or lower motor neuron lesion.

Initial physical therapy assessment was performed by one of three physical therapists blind to treatment group assignment of the subjects. A thorough physical evaluation was carried out prior to treatment and at one month following treatment. In some cases, subjects were excluded from the study if, during the assessment, they did not meet the inclusion criteria or if they presented with exclusion criteria. Subjects were asked to provide background information on their current activity level and an activity rating of high, medium or low was assigned, based on operational definitions. A high activity level was recorded if a subject participated in any recreational or sporting activity for at least 30 minutes, 5 or more time in a week. A medium level of activity was defined as participation for 30 minutes, 3 to 5 times in a week; while a low level was considered any value less than 30 minutes duration or a frequency of less than 3 times a week.

Based on the subjective and objective findings during the evaluation, patients were scored on each item in the PFS (Figure I-1). The scale assigns values to components of the physical and subjective examination, with a cumulative score of 100 representing normal status. In the PFS, Clarke's sign is tested with the knee in extension and the knee in 10 degrees of flexion. A positive sign with the knee in extension is rated with a score of 4, and a positive sign with the knee in flexion is rated with a score of 6. For this study these scores were combined into a score of 10 for a positive test, as the test was only performed with the knee in flexion. The reasons for only performing the test with the knee in flexion are that the test can be quite uncomfortable and the literature does not confirm that there is a difference in findings with the knee in extension versus flexion.

Subjects were given standard instructions on how to record daily pain and function levels using the VAS (Figure I-2) and FIQ (Figure I-3). Subjects were asked to complete the forms at the same time of day for three consecutive days prior to attending their initial treatment session. The VAS was a 10 centimetre horizontal scale that asked subjects to rate their pain at its worst, least and usual levels during the day. The modified FIQ (Chesworth et al., 1989), asked subjects to describe their ability to carry out eight activities that are commonly affected in PFPS. The addition of an "unknown" category was included, as certain functions may not be carried out daily, depending on the individual's activity level.

During the initial assessment a step test was performed, with the patient stepping down from a 6 inch step, leading with the unaffected or least affected leg, and then stepping backwards up the step with the affected or most affected leg. When subjects indicated that they felt knee pain, the test was terminated and the time in seconds to onset of pain was recorded. A standardized step rate was set, using a tape recorded musical beat. The test was discontinued at five minutes if subjects did not complain of knee pain. A second step test was performed when subjects attended their first treatment session, which was normally within two weeks of initial assessment.

Levels of functional limitation were reported by patients, based on an ordinal scale of 5 increments (Figure I-4), at the initial assessment and at the one month reassessment. At the one month reassessment patients also indicated whether their condition had changed. (Figure I-4)

For the purposes of the evaluative study, subjects were randomly assigned to one of three treatment groups which consisted of a home program, supervised exercise program or a program including exercise, patellar taping and biofeedback. One month following entry into the study subjects were reassessed and FIQ and VAS forms were completed over three consecutive days. The step test was only performed once during the reassessment. The goals of this component of the study did not include the effectiveness of different treatment regimes, only the evaluation of the clinical tools used to measure the subjects prior

to any treatment and at a one month follow-up period.

Data Analysis

All data were analysed using the SPSS program on a VAX/VMS computer system Version 5.5. An alpha level of 0.05 was established for al tests.

Consistency

The test-retest reliability of the VAS, FIQ and Step Test were examined prior to treatment, to determine the stability of the scores over a short period of time when real change in the condition should not be expected. Intra-class correlation coefficients (ICC) (Cronbach, 1971) were calculated to determine the level of agreement for the three evaluation scores of the VAS and the two scores of the Step Test. Correlation analysis was performed on the FIQ data. The total index for the FIQ, and the eight questions individually, were analysed for the three sessions. The non-parametric correlation method (Spearman Rho) was used, as the FIQ data was ordinal in nature. Internal consistency of the FIQ and PFS were examined by calculating Cronbach's alpha for each instrument (Kazis, Anderson & Meenan, 1989).

Measure of Clinical Change

Effect sizes were determined for the VAS, FIQ, step test and PFS, using the ratio of the difference of post-treatment and pre-treatment scores by the pooled standard deviation, as defined by Cohen (1988). Patients were sub-categorized as "Improved" if they reported some improvement or significant improvement during their one month reassessment. Those subjects who indicated that their condition was worse or had not changed were classified as "Not Improved". Based on this categorization, effect sizes were determined for the groups (Improved versus Not Improved) for the above mentioned tools.

RESULTS

Demographic and background information on the 56 subjects are presented in Table I-1. The sample population ranged in age from 12 to 41 years of age, with a range of activity levels reported. Pre-treatment and post-treatment scores for the VAS, FIQ and Step Test are described in Table I-2. The mean values for

these tests showed an overall improvement during the one month period. The VAS and FIQ scores are based on three consecutive evaluations for both the pre-treatment and post-treatment evaluations. The step test scores are based on two evaluation sessions at the pre-treatment evaluation and one test at the post-treatment evaluation session.

Intra-class correlation coefficients for the repeated measures of the VAS and step test, presented in Table I-3, indicate a modest level of test-retest reliability for these measures. Spearman Rho Correlation coefficients were calculated over the three days prior to treatment for the individual questions and the combined FIQ. Results presented in Table I-4 represent modest correlation values for most questions. It should be noted that 20 to 30% of patients reported "unknown" for Questions 1 and 7 of the FIQ. With the wide range of activity levels represented in the patient sample, these findings were expected, considering that walking as far as a mile (Question 1) and climbing 4 flights of stairs (Question 7) may not be routinely performed. Both the VAS and FIQ test-retest correlation values improved at the one month re-evaluation.

Descriptive statistics for the PFS and patient self-reporting of functional limitations are presented in Table I-5. These values are based on single test sessions at the pre-treatment evaluation and at the one month reassessment period. Tables II-6 and II-7 present the results of analyses for internal consistency, using Cronbach's alpha model, for the FIQ and PFS. The overall alpha value for the FIQ scores pre-treatment was 0.8463. The post-treatment alpha value for the FIQ was 0.8767. These findings suggest that the eight questions in the FIQ are homogeneous, thus providing evidence that the overall scale is providing more information on function in PFPS than any one item. Question 5, related to problems with sitting, presented with the lowest item-total correlation value.

Cronbach's alpha value for the PFS was 0.6458 . Six items had correlation coefficients of less than 0.2, and when these items were deleted, the overall alpha value increased to 0.7183. Those items with values of less than 0.2 included: use of an orthosis, presence of effusion, degree of patellar crepitus (passive and

resisted), presence of Clarke's sign and limitation of range of motion. The post-treatment alpha value for the total PFS was 0.7720. The lower Cronbach's alpha value suggests that the scale does not strongly measure one factor of the condition. The increased internal consistency of the scale at the one month evaluation period may be explained by the regression of values to the mean on re-evaluation.

Pre-treatment and post-treatment scores for each measure were analysed for the entire sample and two sub-groups. The two sub-groups consisted of those patients who rated their condition as improved after one month of treatment, and those patients who rated their condition as the same or worse after one month of treatment. An analysis of variance comparing the two groups' scores on VAS, FIQ, PFS, step test, and self-report of functional limitations reported no significant differences pre-treatment ($p<0.05$). It should be pointed out that based on the small number of non-improvers ($n=12$) that the power for this test was limited to 0.34, considering a medium effect size. A repeated measures analysis of variance (ANOVA) performed on test scores, showed significant post-treatment differences ($p<0.05$) between the groups for all measures except the step test ($p=0.197$).

Effect sizes for the VAS, FIQ, PFS and step test are presented in Table II-8. Pre-treatment effect sizes were calculated using the first and third day test scores of the repeated testing for the VAS and FIQ, and first session and second session values, in the case of the step test. These effect sizes were calculated in order to describe the changes reflected by each of the tools, prior to treatment, when no or minimal real change should be expected. Both the direction of the change and the degree of change were of interest. The results indicated that for all evaluation methods, small effect sizes, as defined by Cohen (1988), were found during this time period for all tools.

Effect sizes for post-treatment values were determined for the sample population and the two sub-groups of patients, using the first day tests scores for the initial evaluation session and the one month reassessment. The effect sizes calculated for the entire sample, showed that all measures presented with higher

post-treatment effect sizes than pre-treatment values. For those patients indicating an improvement in their condition, effect sizes for the tools were medium to large in magnitude, as defined by Cohen (1988), and positive in direction. For the non-improvers, effect sizes were in the negative direction, except for the worst pain and the step test scores which both showed some improvement. In addition, the effect sizes for the VAS pain scores were small (<0.2), while the FIQ, PFS and step test values were closer to a medium effect size (0.5). As previously mentioned, the repeated measures ANOVA confirmed significant post-treatment differences ($p<0.05$) between the two groups for all tests except the step test.

DISCUSSION

The modest test-retest reliability values for the VAS for pain, the step test and the FIQ can be explained due to the nature of PFPS. As the symptoms and functional limitations are activity dependent, minor changes in pain and function are expected from day to day. Although reliability is a pre-requisite to validity, the degree of reliability is not only dependent on the psychometric properties of the tool, but the stability of the condition. Over a three day period, it is fair to assume that normally the PFPS condition would not change, however, some changes in the symptoms are expected. Therefore the findings of modest intraclass correlation coefficients for each of these tests are reasonable.

An increase in correlation coefficients for the VAS pain scores and the FIQ scores were found when re-test reliability over 3 days was analysed post-treatment. The improved consistency in the scores could be due to: improved stability of the symptoms following one month of treatment; modification of activity level and thus symptoms, due to the treatment intervention; or the Hawthorne effect related to completing the questionnaire. The only exceptions to improved consistency at one month re-evaluation, were the correlation values for Question 5 of the FIQ. The reason for a decrease in these values could possibly be explained by the limited definition for sitting. The other activities in the FIQ are perhaps more consistently performed by an individual, whereas sitting positions can vary greatly. Clarification of the knee position in sitting may improve this

question.

Internal consistency of the FIQ and PFS were evaluated to determine the representativeness of the tools, related to content validity. The FIQ appears to more strongly measure one domain, while the PFS is less homogeneous. The FIQ is designed to evaluate functions related to the patellofemoral joint, therefore it appears to fairly represent the domain of lower extremity function, that is affected in this condition. Future testing of the FIQ must consider whether the functional domain represented is specific to PFPS, or is generalized to other conditions of the lower extremity or knee.

In determining which method of evaluation is most appropriate, the purpose of the tool must be defined. In evaluative studies, it may be better to use a profile index, which represents individual components and does not combine them into a cumulative score. On the other hand, if the purpose of the tool is to classify individuals into groups or disease categories, the cumulative or aggregated index would be more useful.

The PFS combines a number of areas including pain, function and diagnostic tests carried out by the clinician. Indices, such as the PFS, combine a large number of components into a single variable (characterized by a number) which attempts to represent the overall phenomena of interest. Although this method is practical, easy to analyse and attractive in a clinical environment, there are a number of disadvantages. When using the index as an aggregate, information may be lost and detection of clinical change may be more difficult. In addition, interpretation of the score may be difficult if different levels of measurement are used for the individual components, or if certain components are more important than others. Although the PFS appears to have potential as a discriminative indicator of clinical change, as found in this study, reliability testing must be performed prior to validation. In addition, the dimensions of the tool must be more thoroughly evaluated considering content validation.

In using evaluation tools, it is essential to establish reasonable guidelines for determining "what is clinically significant change". Statistical significance is a

necessary condition for proving treatment effectiveness. However, in large sample studies small differences may be found to be statistically significant, when the results are not clinically relevant (Cohen, 1988). On the other hand, what is clinically important may not be statistically significant. The method used in this study, advocated by Kazis et al. (1989), analysed effect sizes for each of the tools over a pre-treatment period and between pre-treatment and post-treatment. Cohen (1988) defines effect size as "the degree to which the phenomenon is present in the population" (pg.9). The effect size index for differences between population means is not the difference between raw difference scores, but the difference between mean "z" standard scores. As a result, effect size represents a dimensionless number, void of raw units. The operational definitions for effect sizes, defined by Cohen (1988), "use levels of effect size which accord with a subjective average of effect sizes such as are encountered in behavioral science." (pg. 13) A limitation that should be kept in mind when interpreting the effect sizes in this study is the smaller sample of non-improvers.

Based on Cohen's guidelines, it appears that effect sizes of 0.2 or less would not be useful indicators of real clinical change for any of the evaluation tools, as small fluctuations were found over the pretreatment period. A small effect size would be the equivalent of the following changes: 0.5 cm. on the 10 cm. VAS for the worst and least pain ratings; one increment on one question in the FIQ; 3 points in the PFS; and 18 seconds in the step test.

For the VAS and FIQ a medium effect size (0.5), as defined by Cohen (1988), would be an appropriate indicator of clinically significant improvement. Again equating a medium effect size to the actual measure, the changes represented would be : 1 cm. on the 10 cm. VAS for the worst and least pain ratings; and either one increment on two questions or two increments on one question in the FIQ. In regards to the non-improved subjects, the findings in this study did not indicate a clinically significant deterioration in pain, considering the effect size. These findings could be explained by the fact that both those individuals who described their condition as the same, and those patients who felt

their condition was worse, were combined in this group. A larger sample of patients who indicated that their condition was worse would have been useful in determining whether the VAS for pain can detect clinical deterioration. Unlike the VAS, the FIQ appears to detect a decrease in the non-improvers, with a medium effect size representing a clinically significant deterioration.

Without pre-treatment effect sizes for the PFS it is more difficult to propose a clinically significant effect size, as the normal day to day fluctuations are not known. Assuming that the pre-treatment fluctuations for the PFS are similar to the other tests, medium effect sizes for both clinical improvement and deterioration may be reasonable estimates. Until the PFS is investigated more thoroughly, a medium effect size, equivalent to 6 points of the total score, can only be proposed as a possible guideline.

In analysing the effect sizes for the post-treatment step test, both the values for the improved and non-improved groups increased in a positive direction. These results suggest that the accuracy of the test is limited relative to group discrimination, although the magnitude of the effect size was greater in the non-improved group. The rationale for the step test procedure was to eccentrically load the affected leg during the descent, which theoretically should increase the patellofemoral joint compression and aggravate the knee (McConnell, 1986). An assumption that was made, based on the literature, was that stair descent created more pain than stair ascent (Bennett & Stauber, 1986). In this study, the majority of subjects reported pain with both climbing and descending stairs. Therefore, the test may not have been as specific as originally intended.

Although stair climbing is affected in PFPS, and was frequently reported in the FIQ as limited, the step test performed in this study did not affectively discriminate between the two groups of subjects. Although the step test was standardized in regards to cadence and step height, only onset of pain was evaluated. As soon as the subject complained of any pain, the test was terminated and the time was recorded. The intensity or duration of pain were not considered in the test, and this could have been an important discriminator

between the groups. To further support the effect size findings, the results of the ANOVA confirm that the step test was the only test that did not demonstrate significant difference between the two groups at the post-treatment period.

CONCLUSIONS

Modest test-retest reliability was found for the VAS, FIQ and step test prior to treatment. These findings were considered acceptable considering that the symptoms associated with PFPS can fluctuate slightly from day to day. High internal consistency was found for the FIQ, which measures functional limitations related to the patellofemoral joint, whereas the PFS appears to measure more than one component of the condition.

Effect sizes were found to facilitate interpretation of the utility of the outcome measures evaluated in this study. In agreement with a previous study (Chesworth et al., 1989) the VAS and FIQ were found to be good discriminators for clinical change. A medium effect size would appear to be a useful guideline for determining clinically significant change when using these tools. Although the PFS discriminated between those patients who improved and those who did not improve, further testing is required to establish levels of reliability and content validity prior to accepting this tool. Finally, the step test, performed in this study, presented with reasonable test-retest reliability. However, the step test was not an accurate discriminative tool.

Figure I-1. Patellofemoral Function Scale

Item 1.	
Pain	
None	10
During vigorous activity	8
During light activity	4
At rest after activity	2
Daily pain irrespective of activity*	0
Item 2.	
Movie sign	4
Absent	0
Present	
Item 3.	
Walking	
No restriction	5
Restricted	0
Item 4.	
Stair climbing	
No restriction	5
Restricted	0
Item 5.	
Jogging	
No restriction	6
Restricted	0
Item 6.	
Sprint & Cutting	
No restriction	10
Restricted	0
Item 7.	
Orthosis**	
None	4
Knee sleeve or shoe insole	2
Total contact knee orthosis	1
Walking cane	0
Item 8.	
Effusion	
None	6
Present	0
Item 9.	
Patellar crepitus (passive)	
None	6
Mild	4
Severe***	0
Item 10.	
Patellar crepitus (resisted)	
None	4
Mild	2
Severe	0
Item 11.	
Quads atrophy	
None	4
>1cm	2
>2cm	1
>3cm	0
Item 12.	
Apprehension Test	6
Negative	0
Positive	
Item 13.	
Clark's sign	
Mild or Negative	10
Severe	0
Item 14.	
Range of motion	
Full	6
<10 limitation	4
>10 limitation	0
Item 15.	
Pain with motion	
Full and pain free	6
Mild pain with resistance	4
Severe pain with resistance	2
Painful no resistance	0
Item 16.	
Squatting	
No problem	8
Slightly impaired	6
Not past 90	2
Unable	0
*Includes night pain	
**Regular use for activity	
***Firm manual resistance compare to other knee	
(Modified from Reid, 1992)	

Figure I–2. Visual Analogue Scale for Pain

The following information is to be recorded at approximately the same time each day (preferably at bedtime). Indicate the severity of your pain today by making a mark on EACH of the three lines below. Mark the point on the line that best indicates your pain level relative to the pain definers at the end of the line.

DATE:

Rate your pain at its worst:

no pain _____ pain as severe as it could be

Rate your pain at its least:

no pain _____ pain as severe as it could be

Rate your pain as it usually felt:

no pain _____ pain as severe as it could be

Figure I-3. Functional Index Questionnaire (FIQ)

The following information is to be recorded at approximately the same time each day (preferably at bedtime). Put a check mark in the column that best describes the way you feel.

Please complete the following:

Today did you have any problem or discomfort in your ___ knee at all with the following activites?

	Unable to Do	Can do with Problem	No Problem	Unknown
1. walking as far as a mile	( )	( )	( )	( )
2. climbing up 2 flights of stairs (16 steps)	( )	( )	( )	( )
3. squatting	( )	( )	( )	( )
4. kneeling	( )	( )	( )	( )
5. sitting for prolonged periods with your knees bent in one position	( )	( )	( )	( )
6. climbing up 4 flights of stairs (32 steps)	( )	( )	( )	( )
7. running a short distance, say 100 meters (about the length of a football field)	( )	( )	( )	( )
8. walking a short distance (about a city block	( )	( )	( )	( )

(Modified from Chesworth et al., 1989)

Figure I–4. Scale for rating functional limitations and change in condition

Functional Limitation:	Change in Condition:
0 None	
1 Annoying	1 Significant improvement noted
2 Limits some activities	2 Some improvement noted
3 Limits most activities	3 No improvement noted
4 Completely disabling	4 Condition worse

Table I-1. Demographics of subjects

Gender	Age (yrs)			Activity Level*		
	Mean	S.D.	Range	Mean	S.D.	Range
Male	25.27	9.94	12-41	2.13	0.92	1-3
Female	24.3	8.1	13-41	1.85	0.83	1-3

*Activity Level rated as 1- Low level, 2- Medium level, 3- High level

Table I-2. Descriptive statistics for VAS Pain, FIQ and Step Test scores pre-treatment and post-treatment

	Pre-Treatment				Post-treatment			
	Mean	S.D.	Range	N	Mean	S.D.	Range	N
VAS*(cm)	2.64	1.66	0.11-6.33	50	1.64	1.58	0-5.33	46
FIQ**	1.27	0.39	0 - 2	51	1.46	0.44	0.38-2	46
STEP TEST***	2.46	1.49	1 - 5	56	3.34	1.7	1 - 5	56

*Worst, least and usual pain ratings for 3 days

**Eight questions over 3 days - based on scale of 0 to 2

***Two tests pre-treatment and one test post-treatment in minutes

Table I-3. Intra-class correlation coefficients for Test-Retest of VAS and Step Test

PRE-TREATMENT

	N	ICC		N	ICC
VAS1	50	0.56	STEP TEST	39	0.63
VAS2	50	0.64			
VAS3	50	0.58			

POST-TREATMENT

	N	ICC	No repeated step test
VAS1	45	0.7	
VAS2	45	0.74	
VAS3	45	0.77	

VAS1 - Worst pain rated on scale of 0 to 10

VAS2 - Least pain rated on scale of 0 to 10

VAS3 - Usual pain rated on scale of 0 to 10

Table I–4. Spearman Rho correlation coefficient values for the total FIQ and individual questions over 3 days prior to treatment

Question	Pre–Treatment Range of Correlations	Post–treatment Range of Correlations
Question 1	0.62 to 0.64	0.67 to 0.80
Question 2	0.62 to 0.74	0.66 to 0.76
Question 3	0.62 to 0.90	0.83 to 0.90
Question 4	0.64 to 0.83	0.87 to 0.93
Question 5	0.83 to 0.87	0.66 to 0.78
Question 6	0.52 to 0.74	0.67 to 0.87
Question 7	0.60 to 0.76	0.79 to 0.88
Question 8	0.43 to 0.53	0.60 to 0.66
Overall	0.69 to 0.77	0.84 to 0.92

Table I-5. Descriptive statistics for Patellofemoral (PF) Function Scale and self report of functional limitation

	Mean	S.D.	Range	N
PF Function Scale*				
Pre-treatment	51.73	13.95	24-88	56
Post-treatment	59.73	16.65	28-92	56
Functional Limitation**				
Pre-treatment	1.84	0.71	0-4	56
Post-treatment	1.5	0.93	0-3	56

*Maximum score of 100 representing normal

**Ordinal scale 0 to 4 (Operational Definitions - Figure I-4)

Table I–6. Internal consistency of FIQ using Cronbach's Alpha Model

	ALL ITEMS ANALYSED	
	Corrected Item–Total Correlation	Alpha If Item Deleted
FIQ1	0.53	0.84
FIQ2	0.78	0.81
FIQ3	0.56	0.83
FIQ4	0.62	0.82
FIQ5	0.40	0.85
FIQ6	0.66	0.82
FIQ7	0.57	0.83
FIQ8	0.58	0.83

ALPHA = 0.8463

Table I-7. Reliability analysis of Patellofemoral Function Scale (PFS) using Cronbach's Alpha Model (n=53)

SCALE ITEMS	ITEM-TOTAL CORRELATION
ITEM 1	0.36
ITEM 2	0.23
ITEM 3	0.39
ITEM 4	0.56
ITEM 5	0.51
ITEM 6	0.52
ITEM 7	-0.21
ITEM 8	0.11
ITEM 9	0.09
ITEM 10	0.14
ITEM 11	0.23
ITEM 12	0.13
ITEM 13	0.17
ITEM 14	0.05
ITEM 15	0.51
ITEM 16	0.22

ALPHA = 0.6458

Table I-8. Size and direction of change for VAS Pain, FIQ, Patellofemoral Function Scale (PFS) and Step Test using mean scores

	All Subjects		Improved		Not Improved	
	Effect	N	Effect	N	Effect	N
VAS						
Pre-treatment (Day1 to Day3)						
Worst Pain	0.02	48	0.04	38	-0.07	10
Usual Pain	-0.2	48	-0.21	38	-0.12	10
Post-treatment						
Worst Pain	0.87	45	1.15	35	0.09	10
Usual Pain	0.43	45	0.75	35	-0.15	10
FIQ						
Pre-treatment (Day1 to Day3)	-0.17	49	-0.17	39	-0.17	10
Post-treatment	0.32	46	0.59	35	-0.5	11
PFS						
Post-treatment	0.63	56	0.81	44	-0.31	12
Step Test						
Pre-treatment (Day1 to Day2)	-0.01	42	0.02	32	-0.16	10
Post-treatment	0.65	56	0.74	44	0.31	12

$$\text{Effect size} = \frac{\text{Post-treatment Scores} - \text{Pre-treatment Scores}}{\text{Pooled Standard Deviation}}$$

Negative sign indicates decreased score

BIBLIOGRAPHY

Abernathy PJ, Townsend PR, Rose RM, Radin EL: Is chondromalacia a separate clinical entity? J Bone Joint Surg [Br] 1978; 60: 205-210.

Bennett JG, Stauber WT: Evaluation and treatment of anterior knee pain using eccentric exercise. Med Sci Sports Ex 1986; 18(5): 526-530.

Bourne MH, Hazel WA, Scott SA, Sim FH: Anterior knee pain. MayoClinProc 1988; 63: 482-491.

Chesworth BM, Culham EG, Tata GE, Peat M: Validation of outcome measures in patients with patellofemoral syndrome. JOSPT 1989; 11: 302-308.

Cohen J: Statistical Power Analysis for the Social Sciences. 2nd ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1988.

Cronbach LJ: Test Validation. In: Thorndike RL, ed. Educational Measurement. 2nd ed. Washington,D.C.: American Council on Education, 1971.

Delitto A: Subjective measures and clinical decision making. Phys Ther 1989; 69: 585-589.

Deyo RA, Patrick DL: Barriers to the use of health status measures in clinical investigation, patient care, and policy research. MedCare 1989; 27(3 Suppl. 1): S254-S268.

Doucette SA, Goble EM: The effect of exercise on patellar tracking in lateral patellar compression syndrome. Am J Sports Med 1992; 20(4): 434-440.

Eisele SA: A precise approach to anterior knee pain: More accurate diagnoses and specific treatment programs. Physician Sportsmed 1991; 19(6): 126-130.

Fairbank JCT, Pynsent PB, van Poortvliet JA, Phillips H: Mechanical factors in the incidence of knee pain in adolescents and young adults. J Bone Joint Surg [Br] 1984; 66B: 685-693.

Ficat RP, Hungerford DS: Disorders of the Patello-Femoral Joint. Baltimore: Williams & Wilkins, 1977.

Fletcher RH, Fletcher SW, Wagner EH: Chapter 3: Diagnosis in : Clinical Epidemiology. 2nd Ed. Baltimore: Williams & Wilkins, 1988: 42-75.

Goldberg B: Chronic anterior knee pain in the adolescent. PediatrAnn 1991; 20:

186-7,190-3.

Insall J: Chondromalacia patella: Patellar malalignment syndrome. OrthopClinNAm 1979; 10: 117-127.

Kazis LE, Anderson JJ, Meenan RF: Effect sizes for interpreting change in health status. MedCare 1989; 3: S178-S189.

Kirshner B, Guyatt G: A methodological framework for assessing health indices. JChronDis 1985; 38: 27-36.

McConnell J: The management of chondromalacia patellae: A long term solution. AustralJPhysiother 1986; 32: 215-223.

McDowell I, Newell C: Measuring Health A Guide to Rating Scales and Questionnaires. New York: Oxford University Press, 1987.

Minns RJ, Birnie AJ, Abernathy PJ: A stress analysis of the patella and how it relates to the patellar articular cartilage lesions. JBiomech 1979; 12: 699-711.

Noyes FR: The Noyes Knee Rating System An assessment of subjective, objective, ligamentous, and functional parameters. Internal Publication of Cincinnati Sportsmedicine Research and Education Foundation. Cincinnati, Ohio (1990).

O'Neill DB, Micheli LJ, Warner JP: Patellofemoral stress. Am J Sports Med 1992; 20(2): 151-156.

Paulos L, Rusche K, Johnson C, Noyes FR: Patellar malalignment: A treatment rationale. Phys Ther 1980; 60: 1624-1632.

Reid DC: Sports Injury Assessment and Rehabilitation. New York: Churchill Livingstone, 1992.

Reikeras D: Patellofemoral characteristics in patients with increased femoral anteversion. SkelRadiol 1992; 21: 311-313.

Saltzman CL, Goulet JA, McClellan RT, Schneider LA, Matthews LS: Results of treatment of displaced patellar fractures by partial patellectomy. J Bone Joint Surg [Am] 1990; 72-A(9): 1270-1284.

Schwarz C, Blazina ME, Siston DJ, Hirsh LC: The results of operative treatment of osteochondritis dissecans of the patella. Am J Sports Med 1988; 16(5): 522-529.

Smythe HA, Helewa A, Goldsmith CH: Selection and combination of outcome measures. JRheumatol 1982; 9: 770-773.

Waisbrod H, Treiman N: Intra-osseous venography in patellofemoral disorders: a preliminary report. J Bone Joint Surg [Br] 1980; 62: 454-456.

Wilson T: Anterior knee pain: A new technique for examination and treatment. Physiotherapy 1990; 76: 371-376.

Wilson JM: Interosseus pressures of the patella in chondromalacia [Dissertation]. University of Alberta, 1985.

CHAPTER THREE

PAPER TWO

DEVELOPMENT OF A CLINICAL TOOL AND PATIENT QUESTIONNAIRE FOR EVALUATION OF PATELLOFEMORAL PAIN SYNDROME PATIENTS

INTRODUCTION

Development of outcome measures has become a priority in the present health care environment. Depending on the health care focus, outcome can mean the result of a disease process of a condition on a person's status; or in evaluative studies, outcome refers to the effects of defined interventions on a patient population. The constant demand to produce evidence of treatment effectiveness has spurred clinical research into validation of outcome measures. In the development of an effective measurement instrument, the researcher must recognize the limitations within the clinical environment. Assessment and treatment time constraints, availability of equipment and facilities and patient compliance and expertise of personnel are some of the many variables that will influence the usefulness of any clinical measurement tool. Unless evaluation methods are practical and easy to use, there will be much difficulty implementing them into the clinical setting.

PFPS (Patellofemoral Pain Syndrome) has been extensively studied in regards to diagnosis and treatment (Reid, 1992). However, little research has been done on the validation of outcome measures. One such study (Chesworth, Culham, Tata & Peat, 1989) reported that a functional index questionnaire (FIQ) and a series of visual analogue scales (VAS) for pain measurement were good indicators of clinical change in PFPS patients. Although other indices composed of clinical tests, functional questions and pain scales have been developed to study efficacy of treatment in patellofemoral disorders, validation studies on these tools have not been reported (Reid, 1992; Noyes, 1990; Saltzman, Goulet, McClellan, Schneider & Matthew, 1990; Schwarz, 1988). Within these indices, the various components are weighted, usually based on clinical judgement of importance, and a cumulative score is calculated.

Linked with the many etiologies of PFPS is the lack of a criterion for measuring change. Factors commonly evaluated in PFPS, primarily from a diagnostic perspective, are extensive, but lack evidence of test sensitivity and specificity. In evaluative studies, diagnostic tools are not always appropriate for

measuring change. Although certain measures, such as medical imaging techniques may be useful in diagnostic testing, their role in monitoring change in PFPS patients is limited.

There are many components to the process of validation. This paper will primarily address the issue of content validity for outcome measures in PFPS. Content validity is established by demonstrating that the components or items in an evaluation method are representative of the domain of interest (Kerlinger, 1973). One of the first stages in the content validation process is to ensure that there are clear definitions for the domain and the dimensions within the domain. The importance of each dimension should be considered, as the weighting of individual components will influence the validity of the tool. Although by definition, content validity depends on subjective judgement alone (Kerlinger, 1973), evidence of other forms of validity will improve content validation.

The objective of the study was to develop an instrument that would be an accurate and responsive measure of change in patients with PFPS. A primary goal was to design a tool and process that would be practical and easily implemented into the clinical setting.

Review of Literature

Based on the review of literature, it was apparent that there are four major areas in evaluation of PFPS: pain, function, activity and clinical testing. Measurement of pain has been extensively researched and a variety of clinical methods for quantifying pain have been published (Fordyce, 1983). Visual analogue scales have been found to be accurate and sensitive methods for measuring pain in patients with knee problems (Flandry, 1991) and PFPS specifically (Chesworth et al. 1989). In functional testing, the functional index questionnaire (Chesworth et al., 1989) appears to be the only published functional index specific to PFPS, which has demonstrated properties of test responsiveness or sensitivity. Activity-related scales have been developed for evaluating knee patients with ligamentous pathologies. Validation of these instruments is ongoing (Noyes, Mooar & Barber, 1990).

Clinical tests frequently reported in the literature represent clinical outcomes ranging from subjective to objective variables. Jette (1989) defined subjective outcomes as those in which the phenomena involve the perception of the patient. Objective outcomes differ in that the phenomena exist independent of the patient's perception. Due to the nature of PFPS, it is obvious that subjective clinical outcomes, such as pain, are essential components of evaluation. Efforts to control for important sources of error commonly encountered in measurement of both subjective and objective outcomes are essential. Objective measurement is one method by which error is reduced. When the term objective measurement is used, the implication is that the bias of the examiner does not affect the results. Many of the evaluation techniques used by clinicians involve judgement, thus introducing tester bias. By standardizing test procedures and more precisely differentiating between normal and abnormal findings, the subjectivity of the measurement can be reduced. Lower extremity alignment is commonly evaluated in PFPS patients to detect biomechanical abnormalities related to femoral anteversion, genu valgum or varum, tibial torsion and pronation of the sub-talar and trans-tarsal joints (Hefzy, Jackson, Saddemi, et al., 1992; Reid, 1992, Goldberg, 1992, McConnell, 1986; Bourne, Hazel, Scott & Simm, 1988). The examiner's observational skills and judgement influence the rating of the degree of abnormality. As large population studies have not been performed to determine normal values for lower extremity alignment, the determination of abnormality and the degree of abnormality are somewhat arbitrary. Lower extremity alignment is influenced by a variety of factors including motor control, joint biomechanics and postural mechanisms. The integration of these factors and the number of structures and regions involved in evaluation make lower extremity alignment difficult to objectively quantify.

Measurement of the Q-angle, representing the orientation of the quadriceps muscle relative to the patella orientation, is frequently reported in the literature (Malek & Magine, 1981; Insall, 1979; Gruber, 1979;Ficat & Hungerford, 1977). Until large population studies are performed to detect true normal estimates, these values should continue to be interpreted with caution. Other measures of patellar

orientation have been proposed, but normal values have been reported based on small clinical populations (DiVeta & Vogelback, 1992; Arno, 1990; McConnell, 1986). Patellar stability, patellar crepitus and knee effusion are further examples of tests relying heavily on the examiner's judgement. Pain with patellar compression (Clarke's sign) and tenderness to palpation of patellofemoral structures are subjective outcomes that are frequently reported in PFPS (Reid, 1992; Goldberg 1992; Insall, 1979, Gruber, 1979). These clinical tests are influenced not only by the patient's perception of pain but also by the examiner's technique. The apprehension test and manual mobilizations of the patella have been used to determine patellar stability (Reid, 1992; Kolowich, Paulos, Rosenberg, et al., 1990). Normal values for patellar movement have not been reported and the sensitivity and specificity of the apprehension test have not been investigated.

Limb girth measurements have been shown to be inaccurate in detecting muscle atrophy in PFPS patients (Doxey, 1987). Muscle performance measured with isokinetic testing has not definitively shown changes in knee extensor or flexor muscle values in PFPS (MacIntyre & Wessel, 1988). Deficits in eccentric torque values of the knee extensors have been reported (Bennet & Stauber, 1986), while Kibler (1987) found deficits in hamstring muscle torque in athletic populations with PFPS. Other studies have concentrated on evaluation of motor control of the muscles associated with the patellofemoral joint using electromyographic techniques. There is some evidence to suggest that PFPS patients may have abnormal activation patterns between vastus medialis obliquus and vastus lateralis muscles, during isotonic leg activities (MacIntye & Robertson, 1992; Voight & Weider, 1991; Hanten & Schulthies, 1990; Souza & Gross, 1990). Although muscle testing using sophisticated instrumentation is considered to be a more objective method of evaluation, the patient's motivation and symptoms will greatly influence muscle activity.

Decrease in flexibility of tensor fasciae latae, hip flexors, hamstrings, rectus femoris and gastrocnemius-soleus muscle groups have been associated with

PFPS, however, the findings are most commonly inferred from the results of treatment versus prospective population studies (Beckman, Craig & Lehman, 1989; McConnell, 1986; Malek & Mangine, 1981). Whether muscle tightness is a cause or an effect of PFPS is not clear. In addition, tightness of these muscle groups is not exclusive to the PFPS patient population and is commonly found in non-symptomatic people.

Temporal components of gait investigated in PFPS patients were found to be insensitive to clinical change (Chesworth et al., 1989). As variations in normal gait are prevalent, more subtle differences, as those found in PFPS patients, may be difficult to detect. In normal gait, the patellofemoral joint reaction forces (approximately 0.5 times the body weight) are much less than in other activities such as stairs and squatting (Woodall & Welsh, 1990). As a result, symptomotology is not commonly reported with walking, except over longer distances. Clinical evaluation of gait is performed over a short distance, with minimal repetition due to time and facility restraints. Abnormal gait patterns may not be evident as symptomotology is not present during the typical clinical gait testing. Although other biomechanical changes such as hyperextension of the knees and excessive pronation of the sub-talar joint are commonly associated with PFPS (Reid, 1992), accurate quantification of these parameters is very limited. In many cases, the presence or absence of a gait component is all that can reasonably be assessed.

Radiological evaluation of PFPS has been shown to be useful in diagnosing pathology at the patellofemoral joint (Reid, 1992). The accuracy and sensitivity of the traditional radiographs in the evaluation of patellofemoral tracking problems continues to be questioned (Shellock, Foo, Deutsch, et al.; 1991, Shellock, 1988; Kujal, Osterman, Kormano & Schlenzka, 1989). Roberts (1989) reported that taping of the patella significantly altered the patellofemoral alignment, measured radiologically, in symptomatic patients. In the same study, however, no correlation was found between radiological measurements and symptoms produced during a step test.

METHODS

The content of the PFPS evaluation tool was developed considering the analysis of a recent evaluative study of PFPS patients and a content validation process using expert reviewers. Twenty-one clinical tests were selected for inclusion in the content validation questionnaire, based on the review of literature.

Evaluative Study of PFPS Patients

An unpublished study on PFPS patients investigated properties of validity for several outcome measures (Harrison, 1994). A sample of 56 PFPS patients were followed over a one month treatment period. In agreement with a study by Chesworth et al. (1989), the findings confirmed that the FIQ and VAS for pain measurement demonstrated properties of test-retest reliability and sensitivity to clinical change in PFPS patients. The FIQ also demonstrated properties of high internal consistency. Certain functional limitations not included in the FIQ were identified, and based on this information, additional questions were developed using the same format as the FIQ.

The ordinal scale used in the study based on patient reporting of duration and frequency of exercise classified activity levels as high, medium or low. As this scale was not found to be sensitive to clinical change, further questions were developed in an attempt to more clearly define activity patterns in this population. The activity questions considered recreational and sport activities, as well as occupation and activities of daily living (Figure II-1).

Content Validation Questionnaire

Based on the first two stages of the study, the content validation questionnaire was developed consisting of three sections (Figure II-2). Section one included visual analogue scales for rating the overall importance of each of the following components of PFPS: pain, disability, activity, subjective clinical tests and objective clinical tests. The second section asked clinicians to rate each of the self-report patient questions as "Needed", "Not Needed" or "Modification Required". These questions covered the areas of functional limitations, pain and activity levels. The 20 clinical tests were included in section three of the

questionnaire. Visual analogue scales were used for rating importance of each test considering the ability of each test to detect clinical change.

Expert Reviewers

It was determined that reviewers should be composed of physical therapists and physicians in different areas of expertise, to improve the utility of the tool. Group One consisted of physical therapists with a recognized expertise in the field of sports physical therapy (Sports Physical Therapists, Sports Physiotherapy Division, Canadian Physiotherapy Association). Group Two consisted of physicians with recognized expertise in sports medicine (Canadian Academy of Sports Medicine Diploma) and orthopaedic surgeons who had clinical experience with PFPS patients. Group Three consisted of physical therapists and physicians who worked primarily in the musculo-skeletal field in acute care or community based settings, but had limited experience with PFPS patients.

Clinicians for each group were identified and recruited for participation in the study. Questionnaires were mailed to the clinicians in January 1993 and collected over a six week period. Data was analyzed on the VAX/VMS computer system version 5.5, using the statistical program SPSSX.

Descriptive statistics were analyzed for all variables. Analysis of variance was performed on the visual analogue data to determine whether any significant differences existed among the three groups on specific variables. Prior to the study, it was determined that a 70% agreement level of all reviewers was necessary for patient questionnaire items to be included.

RESULTS

Responses from 34 out of 36 clinicians (94%) were received and demographics including group assignment, age, years of clinical practice and monthly caseload of PFPS patients are presented in Table II- 1. The three groups were of unequal sample size, with more sports physical therapists being represented in the sample than the other two groups. No significant differences ($p<0.05$) between the three groups were found for the independent variables, except for caseload. Considering that the criteria for inclusion in Group 3 was that

the clinicians did not have extensive expertise with PFPS patients, this finding was anticipated and confirmed that the selection criteria was accurate. The clinicians in Group 2 saw the highest volume of PFPS patients.

Using visual analogue scales, respondents rated the contribution of each of the five evaluation areas (pain, disability, activity, subjective clinical tests and objective clinical tests) relative to their importance in monitoring change in PFPS patients. Considering the mean scores for the three groups, disability was considered to be the most important area followed closely by pain, subjective clinical tests, activity and objective clinical tests. Descriptive statistics for these variables are reported in Table II-2. Correlational analysis indicated a significant correlation ($p < 0.01$) between pain and disability ($r=0.68$), and activity and disability ($r=0.67$). The other correlation values between the five areas ranged from -0.1622 to 0.2840.

The 21 clinical tests were ordered according to the mean scores for all reviewers' ratings of importance in detecting clinical change in PFPS patients. Mean scores from the visual analogue scales for each test are represented in Table II-3. Significant differences ($p < 0.05$) were found between the three groups of reviewers for four clinical tests. Flexibility of hip flexors, flexibility of gastrocnemius-soleus muscle group and knee swelling were considered more important by Group 1, while Group 2 rated radiographs as more important. Analysis of item consistency found Cronbach's alpha value of 0.8352 for the 21 clinical tests, suggesting that subjective and objective clinical methods seem to be representative of one dimension. Based on the high internal consistency for the clinical tests and the inability to define most tests as exclusively subjective or objective clinical outcomes, the area of clinical tests was determined to represent one component of PFPS evaluation.

The questions designed for patient self-rating of pain, activity and function were also critiqued by the reviewers. As stated previously in the methodology section, a question was only considered for inclusion if 70% of the total number of respondents agreed it was needed. Modification of a question was done if 10%

of respondents indicated a change was necessary. All questions met the inclusion criteria of 70% agreement for each of the three groups of expert reviewers. Modifications recommended for 10 questions consisted of redefining scale descriptors or clarifying wording.

PFPS Evaluation Tool

Based on the responses of the expert reviewers, the PFPS evaluation tool was developed consisting of a patient questionnaire and a clinician evaluation form to be completed by the clinician (Figure II-3). The patient questionnaire used a combination of visual analogue scales and categorical measures for rating function, activity and pain. The Lipid Research Clinic Questionnaire (LRCQ) (Figure II-4) (Ainsworth, Jacobs & Leon, 1993) was included as an additional activity measure. This self-report of physical activity has been found to be an adequate indicator of activity over a one month period in a normal population. The tool demonstrated the ability to discriminate between individuals with very low and low physical activity habits, and between those individuals with moderate and high activity patterns. The questionnaire consists of four questions asking individuals to select the most appropriate response. The measure had previously demonstrated evidence of validity and was used as a criterion on which to compare the new activity questions.

The clinical evaluation form included: a medical history, pain rating, activity rating, function rating, lower quadrant scan and clinical tests. Considering the time restraints for assessment, only five clinical tests were selected based on the mean scores for each test. The selected tests with the highest mean scores included: lower extremity alignment, static patellar orientation, passive patellar mobility, flexibility of tensor fasciae latae and flexibility of rectus femoris. Except for the medical history, all components on the evaluation form used visual analogue scales for rating. Operational definitions were developed for each of the clinical tests, as well as for clinicians' ratings of pain, activity and function. As an example, the operational definitions for static patellar alignment are described in Figure II-5.

A pilot study was conducted to evaluate the PFPS evaluation tool. Four

physical therapists with limited expertise with PFPS patients were given the operational definitions for each component and a 15 minute orientation to the form. Each therapist was randomly assigned to evaluate three subjects consecutively. The subjects included two female patients (14 years and 25 years of age) with diagnosed PFPS and one 26 year old male subject with lower extremity mal-alignment, but no patellofemoral joint involvement. Subjects were oriented to the testing process. Each subject completed the patient questionnaire and was asked for comments on the content and clarity of the questions.

Based on the feedback from the therapists using the form, operational definitions were modified for function rating and lower extremity alignment. The activity question asking patients to identify the amount of time that was spent during a day doing specific activities, was reformatted from a visual analogue scale to a categorical scale, based on subjects' feedback. Clinicians found the evaluation form easy to complete within 30 minutes, while patients filled in the patient questionnaire in less than 10 minutes.

DISCUSSION

A content-oriented process is one stage in establishing validity. Content validation regulates the sampling of the construct and ensures that the measurement is comprehensive and representative of the dimensions comprising the domain (Payton, 1988). From the findings of this study, the dimensions of PFPS have been grouped into pain, function, activity level and clinical tests. Kaplan et al. (1976) emphasized that identifying the relative importance of each component of any tool was a critical element of content validity. In this study, relative importance of the dimensions, as rated by reviewers, was not found to be substantially different. The weighting of each area is obviously more important if the tool is to be used as a cumulative index, as the overall score may not reflect important changes within individual components. At this stage of development, the investigators are not convinced that a cumulative index is the most useful method of monitoring clinical change. Using only a combined score of the evaluation areas may result in loss of information relevant to treatment. Furthermore, as the

four dimensions of PFPS are not clearly independent, an additional problem is introduced in combining the dimensions into a cumulative score.

Correlation between disability and pain has previously been reported in PFPS patients (Harrison, 1994). Although intuitively one would consider that pain and activity should also be correlated, based on the ratings of the expert reviewers in this study, low correlation was found. The low correlation between activity and pain has also been reported in individuals with low back pain (Sanders, 1980). People may maintain high levels of activity despite painful conditions. This finding suggests that activity may be a more independent dimension than is commonly considered. As activity level is primarily defined according to intensity, frequency, duration and type of activity, individuals may be able to maintain these parameters, despite painful conditions. For example, a runner with PFPS may maintain his training program, but experience more pain following the activity. In this case, pain and disability ratings may be more appropriate indicators than activity. Individuals may continue to maintain their activity level but have symptoms or problems that result during or following the activity. Noyes (1990) refers to these types of individuals as the "knee abusers" and suggests that measuring change in this population is difficult, as activity level may not alter despite significant pathology and clinical presentation.

Pain

For the purposes of the study, pain was defined as a subjective phenomena based on the individual's perception. Normal fluctuations in pain are present depending on activity, lifestyle, environment and occupation. The questions developed for this component attempted to cover not only rating of pain intensity, but also duration and relationship of pain to functional limitations. Some individuals may be more affected by a low level of pain persistent throughout many activities, while others may be affected more by one or two episodes of intense pain.

Although the current tool has considered pain and disability as separate areas, future testing of the tool may suggest that the two areas can be linked.

One question that must be investigated is whether a better measure can be developed by linking pain, which is inherently a less stable measure in this condition, with disability, which is a more stable measure. Further research is required to address this question.

In order to quantify pain during the clinical evaluation, a standardized test protocol consisting of operational definitions to describe the level of pain during testing was developed. Clinician's rating of pain based on the patient's presentation during the tests were recorded on a visual analogue scale. In order to reduce tester bias, the operational definitions considered only the number of tests or activities in which pain was present and the severity of the pain based on the patient's physical response. There are potentially many factors that can influence a clinician's judgement of a subjective variable such as pain, including: previous clinical experience, cultural influences and personal pain experiences. Although standardization of protocol and rating attempts to reduce measurement error due to tester bias, pain is highly individual and thus establishment of a test that is practical as well as specific for each individual is extremely difficult. Although reliability of a test is a pre-requisite to validity, a reliable test is not necessarily a valid test. In future investigations of PFPS evaluation, one issue that needs to be addressed is whether the clinical rating of pain is related to the patient's rating of pain.

Function

As the PFPS population can vary from sedentary to very active individuals, functional enquiries must encompass a broad range of activities. Although variation in actual performance of any functional position or activity is commonly found in the normal population, it is assumed that individuals, within their own performance, will not significantly differ. The eight questions of the FIQ adequately sampled components of daily living involving lower extremity activities that are commonly affected in PFPS. Additional questions were developed to represent more active lifestyles and occupations, and to further quantify sitting and stair activities.

In goal-oriented outcome research, subjects generate a list of normal functional activities routinely performed that relate to their lifestyle and occupation. This list is used to define the dependent variables on which to measure change (Malec, Smigielski, DePompolo, et al., 1993). This method of determining functional limitations would be very useful in single case design studies, but has some limitations in group designs, as the dependent variables would not necessarily be standardized.

It is a challenge to break down the area of functional outcome measures into workable dimensions. Physical, mental and social dimensions are commonly considered as the key areas in any functional outcome measure. For the purpose of this evaluation tool, primarily the physical dimension of function has been investigated.

Similar to pain measurement, four standardized lower extremity functions including walking, squatting, jumping and stair climbing were included in the clinical evaluation with the clinician rating the performance of these activities using a set of operational definitions and recording the results on a visual analogue scale.

Activity

The set of questions developed to measure activity level were focused more to the active population rather than the sedentary population. The reason for this emphasis was due to the higher prevalence of PFPS reported in active adolescent and young adult populations (Goldberg, 1992). The LRCQ was included as a criterion on which to compare the activity questions. Although limited to four categories of activity rating, the scale can help in detecting the validity of the activity questions. As mentioned previously, the ordinal nature of activity measurement leads to difficulty in quantifying changes. For the patient questionnaire, an activity measure was developed, based on Noyes' work (1990), to quantify the amount of time spent daily performing basic activities requiring various levels of energy and muscle strength. Lower extremity activities requiring varying degrees of knee movement and joint reaction forces at the patellofemoral

joint were selected. Visual analogue scales were initially used for measurement of these activities, in order to increase the responsiveness of the evaluation. However, based on the feedback from the three subjects in the pilot study, a categorical rating system was determined to be more practical. It was determined that although some decrease in responsiveness might result by using a categorical scaling method, that the utility of the tool was more important at this stage of development.

The clinical evaluation component also included a clinician rating of activity based on the patient's history using the operational definitions of activity levels. Again, the clinician rated the activity level using a visual analogue scale.

Clinical Tests

The majority of clinical tests used in evaluation of PFPS patients have not been evaluated for basic properties of reliability, specificity, sensitivity or other forms of validity. The varied etiology of the condition makes the accuracy of testing difficult. As limited information is available on the relationship between certain clinical tests and PFPS, the expert reviewers' results were used as the basis to determine which tests would be included in the tool. The agreement among the three groups of clinicians in regards to the rating of the most important tests for monitoring clinical change, improves the confidence in the utility of these tests. As only a small number of clinical tests can reasonably be included in a clinical tool, those with the highest rating of importance for measuring clinical change were established by the content validation process. Lower extremity alignment, flexibility of tensor fasciae latae and rectus femoris muscle groups, static patellar orientation and patellar mobility are used as the basis for many treatment programs and were rated as the most important tests. When reviewing the clinical test variables, it is important to note that often they cannot be exclusively classified as the cause or the effect of the condition. For example, external rotation of the patella has been linked to a tight tensor fasciae latae muscle group (McConnell, 1986). It is not known whether the tight muscle group develops and promotes mal-alignment of the patella, or whether a change in patellar orientation

leads to the tissue and muscle tightness. The broad etiology of PFPS leads to a large combination of clinical presentations which makes development of a standardized group of clinical tests difficult. Although clinically and theoretically linked, the prevalence of the various clinical findings in PFPS populations has not been reported previously. A goal of the next stage of the study will be investigation of the correlation between the various clinical findings in a PFPS population and a normal population. In addition, the relationship between the clinical tests and the other dimensions of PFPS will be investigated.

As mentioned previously, measurements of pain, function and activity, using visual analogue scales, were incorporated in the clinical evaluation component. Clinicians frequently monitor these variables in determining clinical change. Whether clinician and patient ratings of these three areas are similar must be investigated. Although the components appear to be similar, the influence of the clinical environment, the judgement of the tester (ie. patient and clinician) and the subjective nature of the variables may result in different aspects of these variables being represented in the measures.

CONCLUSIONS

The content validation process in development of a PFPS evaluative tool encompasses many facets. The results of the content validation stage have suggested that the primary areas of evaluation of PFPS patients should include: pain, function, activity level and clinical testing. Standardized evaluation procedures have been defined for the five clinical tests considered by the expert reviewers to be most important in detecting clinical change in PFPS patients. The tests include: lower extremity alignment, patellar orientation, patellar mobility, flexibility of rectus femoris and flexibility of tensor fasciae latae. A patient self-report questionnaire has been developed to measure the areas of pain, function and activity.

Figure II–1. Activity Pattern Questionnaire

16. Considering a typical day during the work or school week check the box which best describes how much time you spend doing each of these activities over 24 hours.

a. SITTING

☐ 7 – 8 hrs./day ☐ 5 – 6 hrs./day ☐ 3 – 4 hrs./day ☐ 1 – 2 hrs./day ☐ less than 1 hr./day

b. LYING

☐ 9 – 10 hrs./day ☐ 7 – 8 hrs./day ☐ 5 – 6 hrs./day ☐ 3 – 4 hrs./day ☐ 1 – 2 hrs./day

c. STANDING

☐ 7 – 8 hrs./day ☐ 5 – 6 hrs./day ☐ 3 – 4 hrs./day ☐ 1 – 2 hrs./day ☐ less than 1 hr./day

d. WALKING

☐ 7 – 8 hrs./day ☐ 5 – 6 hrs./day ☐ 3 – 4 hrs./day ☐ 1 – 2 hrs./day ☐ less than 1 hr./day

e. CLIMBING

☐ 4 flights 2 times/day ☐ 3 flights 2 times/day ☐ 2 flights 2 times/day ☐ 1 flight 2 times/day ☐ 0 flights /day

f. SQUATTING OR KNEELING

☐ More than 16 times/day ☐ 11 – 15 times/day ☐ 6 – 10 times/day ☐ 1 – 5 times/day ☐ 0 times/day

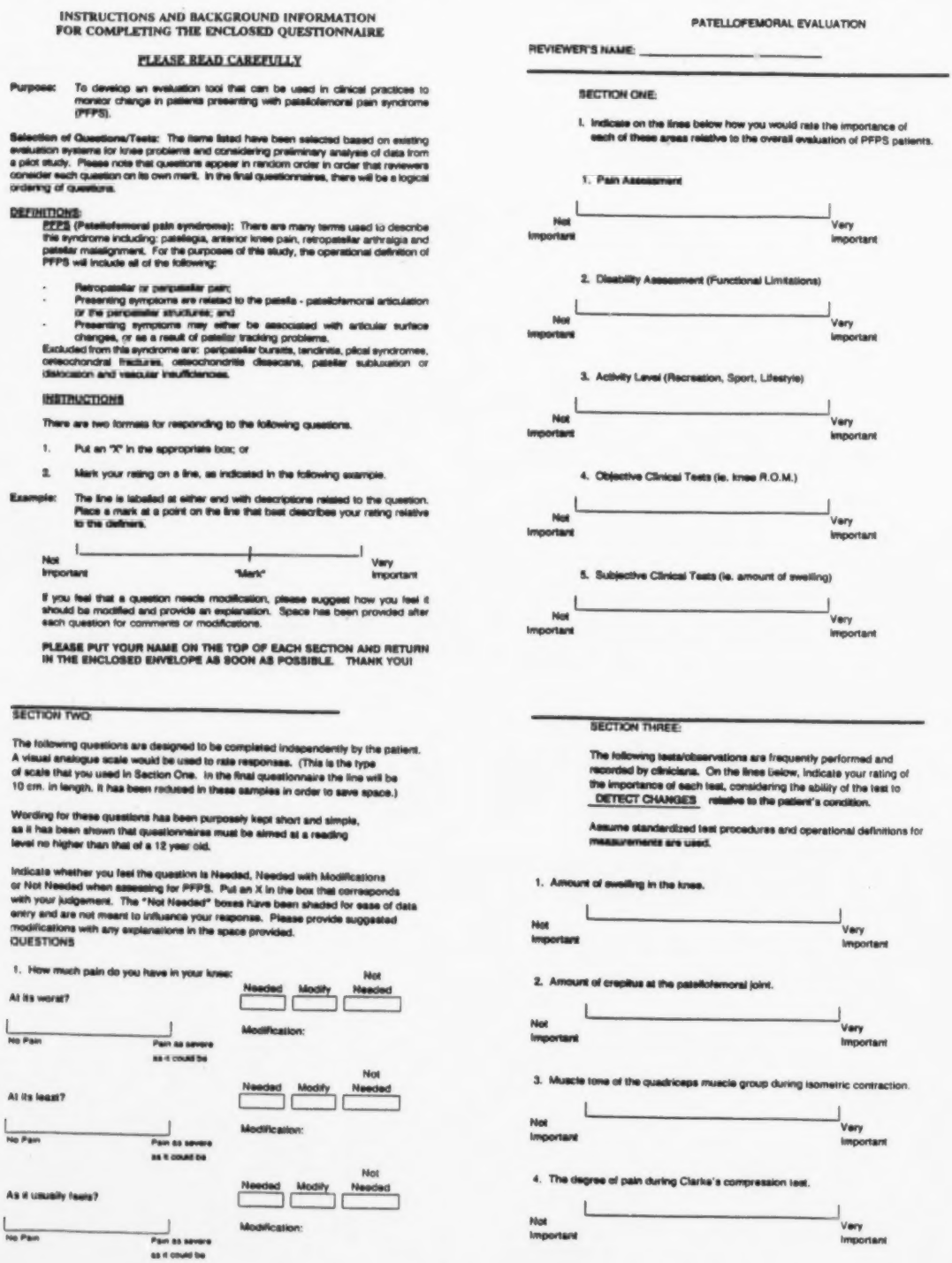
g. CARRYING OR LIFTING

☐ More than 16 times/day ☐ 11 – 15 times/day ☐ 6 – 10 times/day ☐ 1 – 5 times/day ☐ 0 times/day

h. SPORTS OR RECREATIONAL ACTIVITIES

☐ 2 hrs. or more/day ☐ 90 minutes /day ☐ 1 hr. /day ☐ 30 minutes /day ☐ 0 minutes /day

Figure II-2. Content Validation Questionnaire overview



INSTRUCTIONS AND BACKGROUND INFORMATION
FOR COMPLETING THE ENCLOSED QUESTIONNAIRE

PLEASE READ CAREFULLY

Purpose: To develop an evaluation tool that can be used in clinical practices to monitor change in patients presenting with patellofemoral pain syndrome (PFPS).

Selection of Questions/Tests: The items listed have been selected based on existing evaluation systems for knee problems and considering preliminary analysis of data from a pilot study. Please note that questions appear in random order in order that reviewers consider each question on its own merit. In the final questionnaires, there will be a logical ordering of questions.

DEFINITIONS:

PFPS (Patellofemoral pain syndrome): There are many terms used to describe this syndrome including: patellagia, anterior knee pain, retropatellar arthralgia and patellar malalignment. For the purposes of this study, the operational definition of PFPS will include all of the following:

- Retropatellar or peripatellar pain;
- Presenting symptoms are related to the patella - patellofemoral articulation or the peripatellar structures; and
- Presenting symptoms may either be associated with articular surface changes, or as a result of patellar tracking problems.

Excluded from this syndrome are: peripatellar bursitis, tendinitis, plical syndromes, osteochondral fractures, osteochondritis dissecans, patellar subluxation or dislocation and vascular insufficiencies.

INSTRUCTIONS

There are two formats for responding to the following questions.

1. Put an "X" in the appropriate box; or
2. Mark your rating on a line, as indicated in the following example.

Example: The line is labelled at either end with descriptions related to the question. Place a mark at a point on the line that best describes your rating relative to the definers.

Not Important — "Mark" — Very Important

If you feel that a question needs modification, please suggest how you feel it should be modified and provide an explanation. Space has been provided after each question for comments or modifications.

PLEASE PUT YOUR NAME ON THE TOP OF EACH SECTION AND RETURN IN THE ENCLOSED ENVELOPE AS SOON AS POSSIBLE. THANK YOU!

PATELLOFEMORAL EVALUATION

REVIEWER'S NAME: _____

SECTION ONE:

1. Indicate on the lines below how you would rate the importance of each of these areas relative to the overall evaluation of PFPS patients.

1. Pain Assessment

Not Important — Very Important

2. Disability Assessment (Functional Limitations)

Not Important — Very Important

3. Activity Level (Recreation, Sport, Lifestyle)

Not Important — Very Important

4. Objective Clinical Tests (ie. knee R.O.M.)

Not Important — Very Important

5. Subjective Clinical Tests (ie. amount of swelling)

Not Important — Very Important

SECTION TWO:

The following questions are designed to be completed independently by the patient. A visual analogue scale would be used to rate responses. (This is the type of scale that you used in Section One. In the final questionnaire the line will be 10 cm. in length. It has been reduced in these samples in order to save space.)

Wording for these questions has been purposely kept short and simple, as it has been shown that questionnaires must be aimed at a reading level no higher than that of a 12 year old.

Indicate whether you feel the question is Needed, Needed with Modifications or Not Needed when assessing for PFPS. Put an X in the box that corresponds with your judgement. The "Not Needed" boxes have been shaded for ease of data entry and are not meant to influence your response. Please provide suggested modifications with any explanations in the space provided.

QUESTIONS

1. How much pain do you have in your knee:

At its worst? — Needed | Modify | Not Needed — Modification:

No Pain — Pain as severe as it could be

At its least? — Needed | Modify | Not Needed — Modification:

No Pain — Pain as severe as it could be

As it usually feels? — Needed | Modify | Not Needed — Modification:

No Pain — Pain as severe as it could be

SECTION THREE:

The following tests/observations are frequently performed and recorded by clinicians. On the lines below, indicate your rating of the importance of each test, considering the ability of the test to DETECT CHANGES relative to the patient's condition.

Assume standardized test procedures and operational definitions for measurements are used.

1. Amount of swelling in the knee.

Not Important — Very Important

2. Amount of crepitus at the patellofemoral joint.

Not Important — Very Important

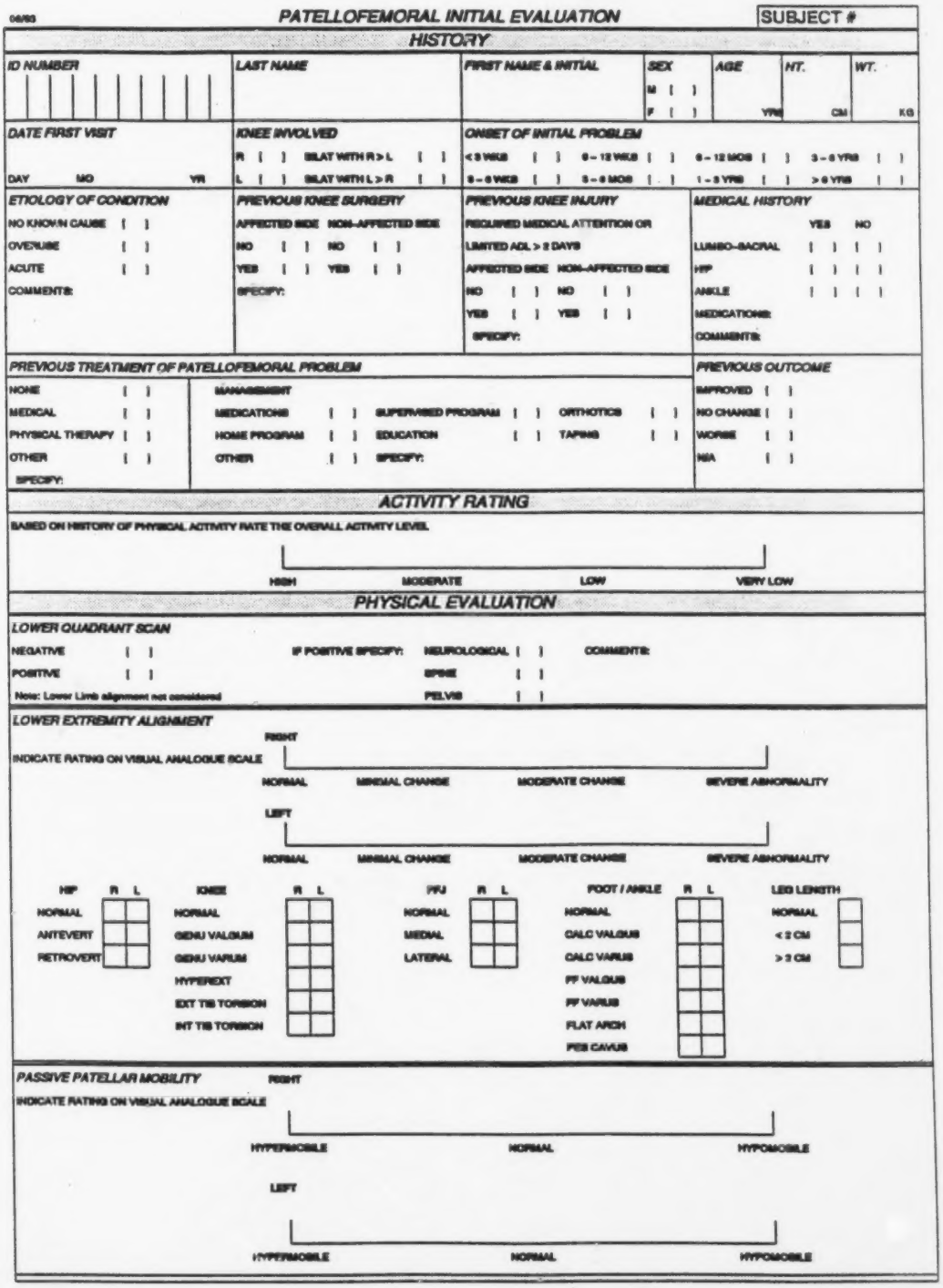
3. Muscle tone of the quadriceps muscle group during isometric contraction.

Not Important — Very Important

4. The degree of pain during Clarke's compression test.

Not Important — Very Important

Figure II-3. Initial Clinical Evaluation Form - Page One



06/93

PATELLOFEMORAL INITIAL EVALUATION

SUBJECT #

HISTORY

ID NUMBER | LAST NAME | FIRST NAME & INITIAL | SEX M [ ] F [ ] | AGE YRS | HT. CM | WT. KG

DATE FIRST VISIT: DAY MO YR

KNEE INVOLVED: R [ ] BILAT WITH R > L [ ] L [ ] BILAT WITH L > R [ ]

ONSET OF INITIAL PROBLEM: < 3 WKS [ ] 3 – 12 WKS [ ] 6 – 12 MOS [ ] 3 – 5 YRS [ ] 3 – 6 WKS [ ] 3 – 6 MOS [ ] 1 – 3 YRS [ ] > 5 YRS [ ]

ETIOLOGY OF CONDITION
NO KNOWN CAUSE [ ]
OVERUSE [ ]
ACUTE [ ]
COMMENTS:

PREVIOUS KNEE SURGERY

AFFECTED SIDE		NON-AFFECTED SIDE	
NO	[ ]	NO	[ ]
YES	[ ]	YES	[ ]

SPECIFY:

PREVIOUS KNEE INJURY
REQUIRED MEDICAL ATTENTION OR LIMITED ADL > 2 DAYS

AFFECTED SIDE		NON-AFFECTED SIDE	
NO	[ ]	NO	[ ]
YES	[ ]	YES	[ ]

SPECIFY:

MEDICAL HISTORY

	YES	NO
LUMBO-SACRAL	[ ]	[ ]
HIP	[ ]	[ ]
ANKLE	[ ]	[ ]

MEDICATIONS:
COMMENTS:

PREVIOUS TREATMENT OF PATELLOFEMORAL PROBLEM

NONE [ ]
MEDICAL [ ]
PHYSICAL THERAPY [ ]
OTHER [ ]
SPECIFY:

MANAGEMENT
MEDICATIONS [ ] SUPERVISED PROGRAM [ ] ORTHOTICS [ ]
HOME PROGRAM [ ] EDUCATION [ ] TAPING [ ]
OTHER [ ] SPECIFY:

PREVIOUS OUTCOME
IMPROVED [ ]
NO CHANGE [ ]
WORSE [ ]
N/A [ ]

ACTIVITY RATING

BASED ON HISTORY OF PHYSICAL ACTIVITY RATE THE OVERALL ACTIVITY LEVEL

HIGH MODERATE LOW VERY LOW

PHYSICAL EVALUATION

LOWER QUADRANT SCAN
NEGATIVE [ ]
POSITIVE [ ]
Note: Lower Limb alignment not considered
IF POSITIVE SPECIFY: NEUROLOGICAL [ ] SPINE [ ] PELVIS [ ]
COMMENTS:

LOWER EXTREMITY ALIGNMENT

INDICATE RATING ON VISUAL ANALOGUE SCALE

RIGHT
NORMAL MINIMAL CHANGE MODERATE CHANGE SEVERE ABNORMALITY

LEFT
NORMAL MINIMAL CHANGE MODERATE CHANGE SEVERE ABNORMALITY

HIP	R	L
NORMAL		
ANTEVERT		
RETROVERT		

KNEE	R	L
NORMAL		
GENU VALGUM		
GENU VARUM		
HYPEREXT		
EXT TIB TORSION		
INT TIB TORSION		

PFJ	R	L
NORMAL		
MEDIAL		
LATERAL		

FOOT / ANKLE	R	L
NORMAL		
CALC VALGUS		
CALC VARUS		
FF VALGUS		
FF VARUS		
FLAT ARCH		
PES CAVUS		

LEG LENGTH	
NORMAL	
< 2 CM	
> 2 CM	

PASSIVE PATELLAR MOBILITY

INDICATE RATING ON VISUAL ANALOGUE SCALE

RIGHT
HYPERMOBILE NORMAL HYPOMOBILE

LEFT
HYPERMOBILE NORMAL HYPOMOBILE

Figure II-3. Initial Clinical Evaluation Form - Page Two

STATIC PATELLAR ORIENTATION

INDICATE RATING ON VISUAL ANALOGUE SCALE

RIGHT

NORMAL — MINIMAL CHANGE — MODERATE CHANGE — SEVERE ABNORMALITY

LEFT

NORMAL — MINIMAL CHANGE — MODERATE CHANGE — SEVERE ABNORMALITY

ROTATION R L
Inferior Border
NORMAL
INTERNAL
EXTERNAL

TILT R L
Medial Border
NORMAL
ANTERIOR
POSTERIOR

SUPERIOR/INFERIOR R L
NORMAL
SUPERIOR (ALTA)
INFERIOR (BAJA)

FLEXIBILITY RECTUS FEMORIS

RIGHT

INDICATE RATING ON VISUAL ANALOGUE SCALE

NORMAL — MINIMAL DECREASE — MODERATE DECREASE — MARKED DECREASE

LEFT

NORMAL — MINIMAL DECREASE — MODERATE DECREASE — MARKED DECREASE

FLEXIBILITY TENSOR FASCIA LATAE

RIGHT

INDICATE RATING ON VISUAL ANALOGUE SCALE

NORMAL — MINIMAL DECREASE — MODERATE DECREASE — MARKED DECREASE

LEFT

NORMAL — MINIMAL DECREASE — MODERATE DECREASE — MARKED DECREASE

PAIN RATING

BASED ON PAIN DURING ASSESSMENT INDICATE AN OVERALL PAIN RATING

NO PAIN — MODERATE — SEVERE PAIN

FUNCTION RATING

BASED ON FUNCTIONAL TESTING INDICATE AN OVERALL FUNCTION RATING

NORMAL — MODERATELY LIMITED — SEVERELY LIMITED

OTHER FINDINGS:

ASSESSMENT COMPLETED ON ONE VISIT YES [ ] NO [ ]

PROGNOSIS

BASED ON ASSESSMENT FINDINGS INDICATE A RATING FOR THIS PATIENT'S PROGNOSIS OVER A ONE MONTH PERIOD

CONDITION WORSE — NO CHANGE — COMPLETE RECOVERY

TREATMENT PLAN:

EVALUATORS NAME _____ PT [ ] MD [ ]

04/93

Circle the best answer for questions 12 to 15.

12. Thinking about the things you do at work or school, how would you rate yourself as to the amount of physical activity you get compared with others of your age and sex?
    a. Much more active
    b. Somewhat more active
    c. About the same
    d. Somewhat less active
    e. Much less active
    f. Not applicable

13. Now, thinking about the things you do outside of work or school, how would you rate yourself as to the amount of physical activity you get compared with others your age and sex?
    a. Much more active
    b. Somewhat more active
    c. About the same
    d. Somewhat less active
    e. Much less active

14. Do you regularly engage in strenuous exercise or hard physical labor?
    a. Yes (Answer Question #15)
    b. No (Stop)

15. Do you exercise or labor at least three times a week?
    a. Yes
    b. No

(Modified from Ainsworth, Jacobs & Leon, 1993)

Figure II-5. Patellar Orientation Operational Definition

Position: Supine with legs relaxed and in normal resting position

Normal Alignment:

Patella mid-line, sitting above the trochlear groove with inferior border in contact with suprapatellar fat pad. The medial border and lateral border of the patella are level (ie. parallel with the frontal plane) and the superior border and inferior border are vertically aligned.

Abnormal Alignment:

Minimal: - Medial border and lateral border of patella are not level

OR

- Superior border and inferior border are not aligned vertically

Moderate: - Any two of the above findings

OR

- Patella baja or alta

Severe: - All components mentioned in minimal

OR

- Patella baja or alta in combination with any of the other findings

Table II–1. Demographics of clinical reviewers

Group Number	Sample Size No.	Age (Yrs.) Mean (S.D.)	Yrs. Practice Mean (S.D.)	PFPS Patient Load/mo. Mean (S.D.)
1	15	37.5 (6.1)	12.7 (4.7)	7 (4)
2	9	41.6 (7.9)	13.4 (8.1)	11 (3)
3	10	41.3 (5.8)	18.6 (5.2)	3 (1)
Total	34	39.7 (6.1)	14.6 (6.3)	7 (4)

Table II–2. VAS rating of importance of each area of PFPS evaluation

Group Number	Disability Mean (S.D.)	Pain Mean (S.D.)	Activity Mean (S.D.)	Subjective Tests Mean (S.D.)	Objective Tests Mean (S.D.)
1	8.1 (1.3)	7.9 (1.7)	6.7 (2.5)	6.5 (2.7)	5.9 (2.6)
2	7.2 (2.9)	6.7 (2.7)	7.1 (2.5)	7.4 (1.9)	6.0 (2.1)
3	7.1 (2.4)	6.7 (2.1)	7.4 (1.4)	7.5 (1.6)	5.7 (1.5)
Total	7.6 (2.1)	7.2 (2.1)	7.0 (2.2)	7.1 (2.2)	5.9 (2.1)

Visual Analogue Scale (VAS) increments = cm.

Table II-3. Ranking of tests to detect clinical change in PFPS patients

Test	Mean (S.D.)
	(mm.)
Lower Extremity Alignment	7.5 (2.1)
Static Patellar Orientation	7.1 (2.4)
Flexibility Rectus Femoris	7.1 (2.4)
Passive Patellar Mobility	6.9 (2.1)
Flexibility of Tensor Fascia Latae	6.9 (2.3)
Gait	6.7 (2.3)
Muscle Tone Quadriceps	6.2 (2.5)
Flexibility Hamstrings	6.2 (2.6)
Flexibility Hip Flexors	6.1 (2.4) *
Knee Swelling	5.8 (2.7) *
Time to Pain with Stepping	5.7 (2.4)
Flexibility Gastrocnemius-Soleus	5.7 (2.6) *
Pain with Passive Patellar Movt.	5.6 (2.6)
Knee R.O.M.	5.2 (2.6)
Flexibility Hip Adductors	4.9 (2.4)
Isokinetic Quadriceps Testing	4.6 (2.8)
Limb Girth	4.6 (3.0)
Clarke's Compression Test	4.3 (2.9)
Crepitus Patellofemoral Joint	4.2 (2.8)
Q-Angle	4.2 (3.0)
Radiographs	3.8 (2.7) *
OVERALL	5.7 (1.2)

* ($p<.05$) between groups

BIBLIOGRAPHY

Ainsworth BE, Jacobs DR, Leon AS: Validity and reliability of self-reported physical activity status: The Lipid Research Clinics questionnaire. Med Sci Sports Ex 1993; 125(No.1): 92-98.

Arno S: The A angle: A quantitative measurement of patella alignment and realignment. JOSPT 1990; 12(6): 237-242.

Barber SD, Noyes FR, Mangine RE, McCloskey JW, Hartman W: Quantitative assessment of functional limitations in normal and anterior cruciate ligament-deficient knees. The Noyes Knee Rating System. An Internal Publication of Cincinnati Sportsmedicine Research and Education Foundation, Cincinnati, Ohio.

Beckman M, Craig R, Lehman RC: Rehabilitation of patellofemoral dysfunction in the athlete. Clin Sports Med 1989; 8: 841-860.

Bennett JG, Stauber WT: Evaluation and treatment of anterior knee pain using eccentric exercise. Med Sci Sports Ex 1986; 18(5): 526-530.

Bourne MH, Hazel WA, Scott SA, Sim FH: Anterior knee pain. MayoClinProc 1988; 63: 482-491.

Chesworth BM, Culham EG, Tata GE, Peat M: Validation of outcome measures in patients with patellofemoral syndrome. JOSPT 1989; 11: 302-308.

DiVeta JA, Vogelback WD: The clinical efficacy of the A-angle in measuring patellar alignment. JOSPT 1992; 16(3): 136-139.

Doxey GE: Girth measurements and muscle bulk. JOSPT 1987; 9(5): 177-183.

Ficat RP, Hungerford DS: Disorders of the Patello-Femoral Joint. Baltimore: Williams & Wilkins, 1977.

Fordyce WE: The Validity of Pain Behavior Measurement. In: Melzack R, ed. Pain Measurement and Assessment. New York: Raven Press, 1983, 145-153.

Flandry F, Hunt JP, Terry GC, Hughston JC: Analysis of subjective knee complaints using visual analogue scales. Am J Sports Med 1991; 19(2): 112-118.

Goldberg B: Chronic anterior knee pain in the adolescent. PediatrAnn 1991; 20: 186-7,190-3.

Gruber MA: The conservative treatment of chondromalacia patella. OrthopClinNAm 1979; 10: 105-115.

Hanten WP, Schultnies SS: Exercise Effect on Electromyographic Activity of the Vastus Medialis Oblique and Vastus Lateralis Muscles. Phys Ther 1990; 70: 561-565.

Harrison EL: Analysis of outcome measures used in the study of patellofemoral pain syndrome. Unpublished manuscript, 1994. (Available from [author's address]).

Hefzy MS, Jackson WT, Saddemi SR, Hsieh Y-F: Effects of tibial rotations on patellar tracking and patello-femoral contact areas. JBiomedEng 1992; 14: 329-343.

Insall J: Chondromalacia patella: Patellar malalignment syndrome. OrthopClinNAm 1979; 10: 117-127.

Jette AM: Measuring subjective clinical outcomes. Phys Ther 1989; 69: 580-584.

Kaplan RM, Bush JW, Berry CC: Health status: types of validity and the index of well-being. Health Serv Res 1976; 11: 478-507.

Kerlinger FN: Foundations of Behavioral Research. 2nd ed. New York: Holt,Rinehart and Winston, 1973.

Kolowich PA, Paulos Le, Rosenberg TD, Farnsworth S: Lateral release of the patella: Indications and cantraindications. Am J Sports Med 1990; 18: 359-365.

Kujala UM, Osterman K, Kormano M, Komu M, Schlenzka D: Patellar motion analyzed by magnetic resonance imaging. Acta Orthop Scand 1989; 60: 13-16.

MacIntyre D, Wessel J: Knee muscle torques in patellofemoral pain syndrome. PhysiotherCan 1988; 40: 20-24.

MacIntyre DL, Robertson DGE: Quadriceps Muscle Activity in Women Runners With and Without Patellofemoral Pain Syndrome. Arch Phys Med Rehabil 1992; 73: 10-14.

Malec JF, Smigielski JS, DePompolo RW, Thompson JM: Outcome evaluation and prediction in a comprehensive-integrated post-acute outpatient brain injury rehabilitation program. Brain Injury 1993; 7:1: 15-29.

Malek MM, Mangine RE: Patellofemoral pain syndromes: a comprehensive and

conservative approach. JOSPT 1981; 2: 108-116.

McConnell J: The management of chondromalacia patellae: A long term solution. AustralJPhysiother 1986; 32: 215-223.

Noyes FR: The Noyes Knee Rating System An assessment of subjective, objective, ligamentous, and functional parameters. Internal Publication of Cincinnati Sportsmedicine Research and Education Foundation. Cincinnati, Ohio (1990).

Payton OD: Research: The Validation of Clinical Practice. 2nd ed. Philadelphia: F.A.Davis Co., 1988.

Reid DC: Sports Injury Assessment and Rehabilitation. New York: Churchill Linvingstone, 1992.

Robert JM: The effect of taping of patellofemoral alignment. A radiological pilot study. In: Proceedings of the Sixth Biennial Conference of Manipulative Therapists Association of Australia, Melbourne, 1989; 146-151.

Saltzman CL, Goulet JA, McClellan RT, Schneider LA, Matthews LS: Results of treatment of displaced patellar fractures by partial patellectomy. J Bone Joint Surg [Am] 1990; 72-A(9): 1270-1284.

Sanders SH: Toward a practical instrument system for the automatic-measurement of up-time in chronic pain patients. Pain 1980; 1: 103-109.

Schwarz C, Blazina ME, Siston DJ, Hirsh LC: The results of operative treatment of osteochondritis dissecans of the patella. Am J Sports Med 1988; 16(5): 522-529.

Shellock FG, Foo TKF, Deutsch AL, Mink JH: Patello-femoral joint: Evaluation during active flexion with ultrafast spoiled GRASS MR imaging. Radiol 1991; 180: 581-585.

Shellock FG, Mink JH, Fox JM: Patellofemoral joint: kinematic MR imaging to assess tracking abnormalities. Radiol 1988; 168: 551-553.

Souza DR, Gross MT: Comparison of vastus medialis obliquus: vastus lateralis muscle integrated electromyographic ratios between healthy subjects and patients with patellofemoral pain. Phys Ther 1991; 71: 310-316.

Voight ML, Wieder DL: Comparative reflex response times of vastus medialis obliquus and vastus lateralis in normal subjects and subjects with extensor mechanism dysfunction. Am J Sports Med 1991; 19(2): 131-137.

Woodall W, Welsh J: A biomechanical basis for rehabilitation programs involving the patellofemoral joint. JOSPT 1990; 11(11): 535-542

CHAPTER FOUR

PAPER THREE

VALIDATION OF A CLINICAL EVALUATION TOOL

FOR PATELLOFEMORAL PAIN SYNDROME

Introduction

Evaluation of clinical change is an important component of health care delivery. Without valid methods of determining change in a patient's status, appropriate treatment is jeopardized. In order to effectively determine rationalization of health care services, research into outcome measures is essential. Tools must demonstrate properties of validity if they are to be used in making decisions on priority of care and effectiveness of medical intervention.

Similar to many conditions, investigation of the efficacy of treatment techniques for PFPS (Patellofemoral Pain Syndrome) has infrequently addressed the issue of validity of the evaluation methods (Chesworth, Culham, Tata & Peat, 1989). In order to be confident with the results of clinical studies, measurement methods must demonstrate appropriate levels of validity. Validation of a measurement system or method consists of many stages and is not a finite process. Measurements must at least be able to discriminate between those populations with the condition and those without (Payton, 1988). Considering that the etiology of PFPS is so nebulous and that the clinical signs and symptoms associated with the condition may be present in non-PFPS populations (Noyes, 1990), evaluation methods must be investigated for discriminant validity.

Evidence of construct validity is the ultimate goal in development of a measurement method (Kaplan, Bush, & Berry, 1976). Information on construct validity can be gathered from determining the relationship between the new measures and existing ones. In the case of conditions like PFPS, previously validated measures are extremely limited. Therefore, development of measurement tools requires the clinician to determine reasonable hypotheses based on current theory and clinically relevant findings (Travers, 1951). One method that has been traditionally used to determine construct validity is the multi-trait multi-method technique (MTMM) developed by Campbell and Fiske (1959). Following the determination of the basic components of a condition through a process such as content validation, two independent methods of measuring each component are identified. The MTMM considers the relationships between the various variables

and methods, and determines levels of construct validity based on specific criteria (Ferketich, Figuerdo & Knapp, 1991). The basis for construct validation using this method is that a variable should correlate higher with the same variable, measured with a different technique, than with other variables even if the method is the same (Campbell & Fiske, 1959).

Investigation of the properties of sensitivity and specificity of evaluation tools relates to diagnostic discrimination of tests. Sensitivity measures the proportion of those individuals who present with positive tests when they have the condition or disease. Specificity measures the proportion of those individuals without the disease who appropriately test negative (Fletcher, Fletcher & Wagner, 1988). Obviously, lack of sensitivity or specificity will greatly reduce the validity of the test. Lower extremity alignment is an example of a test that may be positive in individuals without PFPS (Noyes, 1991). Determining the sensitivity and specificity for clinical tests are essential steps in building valid tools of assessment.

The purpose of this study was to analyse psychometric properties of several components of a clinical outcome measure for evaluating PFPS patients. As validity is an ongoing process, the goals of the study were to present information on the strengths and weaknesses of the evaluation methods and to determine reasonable guidelines for using the measures in the clinical settings.

Methodology

The tools that were evaluated in this study were previously developed using a content validation process. The tools consisted of a clinical evaluation form and a patient questionnaire (Harrison, 1994). Operational definitions and standardized protocols were established for rating each of the clinical tests: pain, function, activity level, and bilateral lower extremity alignment, patellar orientation and patellar mobility and bilateral flexibility of rectus femoris and tensor fasciae latae muscle groups. Ten centimetre visual analogue scales were used for measuring each of these variables.

Lower extremity alignment was evaluated in the standing position with the therapist rating the alignment from normal to severe abnormality. Eight areas

considered in the evaluation included: the femoral neck to shaft relationship, the tibio-femoral angle in the frontal plane, tibial torsion, knee hyperextension or flexion, patellofemoral position in relation to the foot, tibio-calcaneal angle, forefoot angle, and medial longitudinal arch position (Magee, 1992; Hefzy, Jackson, Saddemi & Hsieh, 1992; Reid, 1992; Donatelli, 1990; McConnell, 1986). Leg length was also observed in standing. If less than three of the eight components demonstrated abnormality and the abnormality corrected with minimal voluntary effort from the subject, it was rated as minimal abnormality. A moderate abnormality was described if less than three of the components were abnormal and correction required much voluntary effort from the patient or external support; if three to five components demonstrated abnormality and corrected with minimal voluntary effort; or leg length discrepancy was greater than two centimetres. If five to eight of the components were abnormal the lower extremity alignment was rated as severely abnormal.

Passive patellar mobility was rated as hypermobile to hypomobile based on the passive movements of the patella with the subject positioned in supine. Lateral, medial, superior and inferior patellar movements were performed by the therapist. Patellar movement and tissue tension were considered in the rating (Reid, 1992). Patellar orientation was also assessed with the subject in supine lying and the rating ranged from normal to severe abnormality. A normal rating was given if the patella was mid-line, sitting above the trochlear groove with the inferior border in contact with the suprapatellar fat pad, and if the lateral and medial patellar borders were level in the frontal plane (Reid, 1992; McConnell, 1986). Minimal abnormality was described if one of the following patellar positions were noted: the medial and lateral borders were not level, the superior border and inferior borders were not aligned vertically or the patella was sitting superior or inferior to the normal position by less than one centimetre. A moderate abnormality was described if two of the above patellar positions were found or the patella was greater than one centimetre superior or inferior to the normal position. Severe abnormality was described if all the previous components were found or

the patella was more than one centimetre superior or inferior relative to the normal position and there was at least one other abnormal position.

Flexibility of the rectus femoris and tensor fasciae latae muscle groups were tested with the subject in the Thomas test position and ratings were given for each muscle group ranging from normal to marked decrease (Magee, 1992; Kendall & Kendall McCreary, 1983). For the rectus femoris muscle group, normal flexibility was defined if the test hip could obtain a neutral position in the sagittal plane, with the knee flexed to 90 degrees. Specific knee flexion range of motions were defined for each of the ratings: minimal abnormality (70 to 90 degrees), moderate abnormality (50 to 70 degrees) and severe abnormality (less than 50 degrees). Normal flexibility of tensor fascia latae was defined if the test hip was positioned in the neutral considering the frontal plane, with no pelvic or lumbar spine compensation. Minimal abnormality was described if the hip was held in less than 10 degrees abduction or less than 10 degrees of internal rotation, and if the therapist could passively position the hip into a neutral position without pelvic or lumbar spine compensation. Moderate abnormality was described if either of the positions described above were greater than 10 degrees and if the therapist was unable to position the hip in neutral without pelvic or lumbar spine compensation. If hip abduction and internal rotation were greater than 10 degrees in the resting position and the therapist was unable to passively position the hip in a neutral position this was considered a severe abnormality.

Evaluation of activity by the clinician was rated from very high to very low, considering the patient's history of current frequency and intensity of physical activity (Ainsworth, Jacobs & Leon, 1993; Noyes, 1990). A very high rating was given if the subject participated in intense physical activity five to seven times per week for a minimum of one hour. Moderate activity was described if the subject participated in activity three to five times per week for a minimum of 45 minutes. Very low activity was defined if the subject did not participate in physical activity at work, school or outside of work or school. The clinician's functional rating was based on the subject's ability to perform six activities in the clinical setting. The

activities included: walking a short distance, squatting, jumping, kneeling, stepping up and stepping down. Overall performance of these activities was rated from normal to severely limited. Normal was defined if the subject performed all activities without pain and with no observable compensatory movements. If the subject had pain or compensation with three activities or could not complete one activity but could complete the others with pain in one or more of the activities, this was defined as moderately limited. Severely limited was described if the subject could not complete three or more of the six activities, or demonstrated severe compensatory movements in four or more activities.

Clinicians were oriented to the protocols and standardized recording methods. In addition, an instructional video demonstrating the use of the form and evaluation techniques was reviewed by therapists. Six physical therapists volunteered to participate as evaluators. Four of the therapists worked in private practice settings, while two of the physical therapists worked in acute care settings. Patients referred to the clinical facilities over a six month period, meeting the inclusion and exclusion criteria indicated in Table III-1, were asked to volunteer for the study. Information on the study was given to each patient and informed consents were signed. In order to control for growth and development factors and to ensure a reasonable representation of the various age groups, patients were grouped by age (12 to 19 years; 20 to 27 years; 28 years and older) (Nordin & Frankel, 1989).

Questionnaire items, based on a previous content validation study (Harrison, 1994) covered the areas of pain, function and activity. Six pain questions, 13 function questions and three activity questions were rated independently by the patient using visual analogue scales. The LRCQ (Lipid Research Clinic Questionnaire), consisting of four questions using a categorical scale, was included in the questionnaire. The LRCQ has been previously used to classify healthy populations as highly active, moderately active, low active and very low active, depending on the combination of responses (Ainsworth, Jacobs & Leon, 1993). In addition, evaluation of activity level also consisted of investigation of

activity patterns. Subjects rated their own frequency of participation over a 24 hour period in eight activities including: standing, sitting, lying, walking, climbing, squatting or kneeling, carrying or lifting, and sports or recreational activities. These activities were based on previous work done by Noyes (1990) and in consideration of the most common activities affected with PFPS (Reid, 1992; Fulkerson, 1990; Chesworth et al., 1989). Five increments for each of the eight activities resulted in a cumulative ordinal scale with a range of 0 (very low) to 40 (very high).

Subjects completed the questionnaries prior to being evaluated by the physical therapist, in an attempt to keep the clinicians' ratings independent from the subjects' ratings. The physical therapists did not review the results of the subjects' questionnaires. Treatment programs were individualized for each patient. A variety of treatment techniques and programs, including exercises, taping, modalities and orthotics were given to patients. Some patients were instructed in home programs, while other patients were seen regularly over the month period. The treatment options were determined by the therapist in conjunction with the patient. Testing was repeated after a one month period or if the patient was discontinued from treatment prior to one month re-evaluation was performed at this time.

A second group of subjects from 12 to 27 years of age, without PFPS were recruited from the University of Saskatchewan staff, student body and summer sport camps. These subjects were matched by age with the PFPS group. The non PFPS subjects were evaluated twice over a one week period, by one of the six physical therapists who assessed the PFPS patients. The same evaluation tools and protocols were used with this group of subjects. One item was eliminated from the questionnaire as it referred specifically to change in the knee due to PFPS.

Data Analysis

Visual analogue scales, measured in centimetres, and ordinal and categorical data were entered into the relational database Paradox 3.5 (Borland,

Inc.). Data was analysed using the statistical package SPSSX on the VAX/VMS computer system. For all clinical tests and questionnaire items on function, pain and activity that used visual analogue scales, parametric analyses were performed. Additional activity questions consisted of both categorical and ordinal scales, therefore, non-parametric analyses were used for these data. A significance level of 0.05 was established for all tests.

Descriptive statistics were analysed for all subjects. Internal consistency was investigated for: pain questions, function questions, activity questions and five clinical tests, by calculating Cronbach's alpha coefficients (Kazis, Anderson & Meenan, 1989) for the four individual areas, as well as some of the components combined. Intraclass correlation coefficients (ICC) (Cronbach, 1971) were calculated on the four areas to determine test-retest stability between the two test periods on the non-PFPS group, as these subjects were assumed to represent a stable population in regards to the dependent variables. A repeated measures ANOVA (analysis of variance) was carried out between age matched PFPS and non-PFPS groups in order to investigate one aspect of discriminant validity. A multi-trait multi-method (MTMM) analysis was performed using the PFPS group data, in order to investigate construct validity (Campbell & Fiske, 1959). For this study, two methods (patient report through the questionnaire and clinician evaluation) and three traits (pain, function, activity) were examined. In addition, a principal component factor analysis with varimax rotation was performed on the same data to further evaluate construct validity (Figueredo, Ferketich & Knapp, 1991). The goal of principal component analysis is to explain the maximal variance in a set of variables, on the basis of a few underlying constructs. The analysis focuses on explaining the total variance in the observed variables. In comparison, factor analysis uses a statistical model that partitions the total variance into common and unique variance and focuses on explaining the common variance in the observed variables. Principal components analysis is uses to determine the dimensions of the factors and to detect outlying or unusual observations (Dunteman, 1989). Each factor has associated with it an eigenvalue,

which is the amount of variance in the data explained by the factor. Only those factors with eigenvalues greater than one were extracted, as these factors account for the majority of the variance. A factor loading matrix for the six variables was analysed and each variable was assigned to the factor on which it loaded most highly.

Levels of criterion validity for the activity questionnaire items were investigated through correlational analyses between the scores on the LRCQ, which had been previously validated on a healthy population (Ainsworth, Jacob & Leon, 1993), and the activity questions designed for the study.

The five clinician rated tests including lower extremity alignment, patellar mobility, patellar orientation, flexibility of rectus femoris and flexibility of tensor fasciae latae were evaluated for sensitivity and specificity (Fletcher, et al., 1988). The gold standard used to define diseased versus non-diseased individuals has previously been described in the inclusion and exclusion criteria (Table III-1). For determining sensitivity and specificity of clinical methods, the tests must be classified dichotomously. It was necessary to determine cutoff values for positive versus negative tests on each of the visual analogue scales for the five tests. Based on the operational definitions and the clinical literature, a rating of minimal involvement or less than minimal involvement, recorded on the visual analogue scale, was considered to be a negative result, while anything greater than minimal involvement was considered as a positive test. As the patellar mobility scale ranged from 0 to 10 cm., with 5 cm. indicating normal, a range of 4 to 6 cm. was considered within normal range, while 0 to 3.9 cm. and 6.1 to 10 cm. were considered abnormal. Two by two tables were constructed using the gold standard for PFPS and non-PFPS (diseased and non-diseased) as one category, and the frequency of the positive and negative findings for each test as the second category.

Effect sizes were determined for all variables using the ratio of the difference between initial and final test mean scores to the pooled standard deviation (Cohen, 1988). The major goal of analysing effect sizes was to establish

guidelines for determining clinically relevant change for the various variables and components of the evaluation tool.

Results and Discussion

Forty-one patients were recruited for the study, and 28 of these subjects were matched by age with non-PFPS subjects. Demographics for the 41 PFPS subjects, the 28 matched PFPS subjects and the 28 non-PFPS subjects are presented in Table III-2. The total patient population reported a majority of bilateral knee problems (61%) with onsets of the condition ranging from two weeks to more than six years. PFPS conditions were categorized as due to the following etiologies: overuse (48%), trauma (15%) or no known cause (37%). Descriptive statistics for the clinical variables and the questionnaire items are indicated in Tables III-3, III-4A and III-4B.

The results of the reliability analysis for internal consistency of the various components are presented in Table III-5A. The questionnaire component for function was highly homogeneous, while clinical tests and pain were found to be lower. With only three activity questions using visual analogue format, very low correlation values were found. Pain and function questions combined, resulted in higher correlation values, providing evidence to support the homogeneity of the sample. Intra-class correlation values for the non-PFPS subjects are presented in Table III-5B. These results represent test-retest values on subjects who were not expected to change during this time frame. The combined clinical tests demonstrated very high values, while modest values were found for the questionnaire components.

For all variables, patient ratings revealed more limitations than non-patient scores. The results of the ANOVA, using matched subject data, demonstrated significant differences between the PFPS population and the non-affected population for all clinician measured variables except patellar mobility ($f=0.876$). Also, significant differences ($p<0.05$) were found between the groups for questionnaire ratings of pain and function, but not for questionnaire ratings of activity ($f=0.152$). Based on Friedman's analysis of variance for non-parametric

data, the ordinal questionnaire items that did not show significant difference between the groups at the 0.05 level were the LRCQ ($p=0.23$) and the Activity Pattern score ($p=0.54$). Between group analysis of the individual activity pattern questions revealed that four of the eight questions involving: standing, walking, squatting/kneeling, and sports/recreation, were significantly different. No significant differences were found between the other activities: sitting, lying, climbing and carrying/lifting. Considering a large effect size, the power for detecting significant difference with these tests was 0.9 (Cohen, 1988).

The results of the MTMM, presented in Table III-6, showed evidence of convergent validity due to significant correlation values between each of the three variables (activity, pain and function) measured with the two methods of clinician rating and patient rating. Correlation values between activity rated by the two methods were higher than the correlation values between activity and function and activity and pain. The clinician rating of activity correlated significantly with clinician ratings of pain and function, whereas the patient rating of activity did not correlate significantly with the patient ratings of pain and function. These results suggested that activity was distinct from pain and function, when considering the subject questionnaire scores. The finding of high intercorrelation values (0.43, 0.43 and 0.71) between the three variables within the clinician rating method indicated that there was shared method variance. The correlation matrix further revealed that pain and function were correlated higher than when either of the variables were correlated with each other using the second method. These results indicated that pain and function, as measured in this study, were not separate dimensions of PFPS.

Principal component factor analysis with varimax rotation further confirmed that only two factors were represented considering the six methods. (Table III-6B) High loadings were found on factor one for function and pain rated by both the subject and clinician, while activity variables rated by the subject and clinician loaded highly on the second factor. The principal component analysis confirmed that activity represents a separate component. Function and pain measures

appeared to share common variance, therefore, they should not be considered as separate dimensions within PFPS evaluation.

Correlational analyses of the activity questionnaire items investigated the relationship between the LRCQ, as the criterion, the visual analogue rating of frequency of participation, the combined visual analogue ratings of three activity questions, and ordinal ratings of activity pattern. Although all correlations were significant at the 0.05 level, the values were low (0.24 to 0.47). When the two groups were individually analyzed, results indicated significant correlation values in the PFPS group for all activity methods ranging from 0.51 to 0.69, except for the combined visual analogue ratings of the three activity questions. Considering the non-PFPS group, no significant correlation values, between the three methods and the LRCQ, were found with values ranging from -0.1 to 0.27. These results suggested a modest relationship between the various questionnaire methods of measuring activity in the patient population, except for the cumulative rating of the visual analogue scales for three activity questions. The findings suggested that frequency of participation and four of the activity pattern questions were the better methods of categorizing patients' activity levels.

The sensitivity, specificity and positive and negative predictive values for the clinician rated lower extremity musculo-skeletal tests are presented in Table III-7. When interpreting these findings, it must be remembered that cutoff points on each visual analogue scale for determining positive and negative findings were based on the review of literature and operational definitions for each test. The results revealed that the probability of tests being positive with the PFPS population ranged from very limited, in the case of patellar orientation, to fair, considering patellar mobility. All tests demonstrated reasonable levels of specificity, suggesting that the clinical methods were successful at detecting negative findings in the non-patient population. Higher positive predictive values indicated that if a test was positive, there was a good probability that the individual had PFPS. The modest negative predictive values revealed that negative findings were not as successful at distinguishing between the PFPS and non-patient

groups, except for patellar mobility. Frequency analysis of the five clinical tests demonstrated the following number of positive tests for each of the 41 PFPS subjects: 4 % with none, 17% with one, 31% with two, 10% with three, 17% with four, and 21% with five. When considering the non-PFPS group, the breakdown showed that 85% of the subjects did not present with any positive tests, while 10% presented with one positive clinical test and 3% presented with two positive clinical tests.

Based on the ratio of the difference between initial and final mean scores to the pooled standard deviation, effect sizes for the clinician rated variables and questionnaire items for the PFPS and non-PFPS groups were calculated (Table III-8). As the non-PFPS group was considered to represent a stable population, it was proposed that the scores should not change over a short period of time. Therefore, an important stage in determining clinically relevant change in a patient population was to establish effect sizes for the tools on a non-PFPS group. The results indicated that a small effect size would not be considered clinically relevant for clinician rating of lower extremity alignment, flexibility of rectus femoris, flexibility of tensor fasciae latae, activity level, and subject rating of activity and pain. In the non-PFPS group, closer to medium effect sizes (d=0.5) were found for clinician ratings of patellar mobility, patellar orientation, and pain; and subject rating of function. Therefore, a medium effect size could not be considered clinically relevant for these tests. Keeping in mind that there is no criterion for determining actual change in the PFPS group from initial test session to post treatment test session, effect sizes must be cautiously interpreted for the patient sample. Large effect sizes (d=0.8) were found for clinician ratings of function, pain, and flexibility of rectus femoris and tensor fasciae latae muscle groups; and for patient rating of pain. Medium effect sizes (d=0.5) were represented for clinician rating of lower extremity alignment and patient ratings for one activity measure and function. Patellar orientation values were closer to a medium effect size (d=0.37), while patellar mobility, the clinician ratings of activity and patient ratings of two of the activity measures represented a small effect size.

The effect sizes established for cumulative scores considering clinician ratings and questionnaire items are presented in Table III-9. Activity ratings have not been included in the cumulative scores. The findings suggested that a medium effect size may be a useful indicator of clinically relevant change in PFPS patients if using the combined measures. At the least, the results indicated that effect sizes of 0.2 to 0.3 are not clinically relevant as these values are found in a non-PFPS population.

Patients in this study were representative of the PFPS populations commonly reported in the literature (Reid, 1992; Goldberg, 1991; McConnell, 1986; & Fairbank, Pynsent, van Portvliet & Phillips, 1984). A young adult, primarily female population was assessed. Patients presented with a majority of bilateral knee problems and overuse etiology. Both the PFPS and non-PFPS groups were physically active, based on the results of the activity classification. Goldberg (1991) reported a higher prevalance of PFPS in active adolescent and young adult populations, thus supporting the findings of high activity levels in the PFPS subjects. As the exact etiology of the condition is not clear, increased activity may be one of many factors related to PFPS. High levels of activity may account for exposure to risk factors associated with the condition. In addition, the active individual may present with more symptoms and limitations from the condition, due to the demands of a physically active lifestyle.

The assessment of PFPS patients has traditionally consisted of a number of clinical tests of the patellofemoral joint and lower extremity (Hefzy, et al., 1992; Goldberg, 1991, McConnell, 1986; Insall, 1979). Clinicians tend to rely on what they consider to be more objective clinical tests than patient self-reports (Jette, 1989). The findings of this study suggest that there are some limitations and merits to both types of evaluation.

Many of the clinical tests used in the diagnosis and treatment of musculo-skeletal problems are affected by tester bias. In this study, an attempt was made to decrease error introduced by tester bias by standardizing protocols and developing operational definitions for each test. The high ICC values found for test-

retest consistency in the non-PFPS subjects provides support for the reliability of the clinical tests.

Previously, patient questionnaires have been used to quantify limitations in function and levels of pain in PFPS patients and have proved to be useful (Chesworth et al, 1989). With proper attention to questionnaire design and scaling, self-reports on variables, especially those involving patient's perceptions, can tap into dimensions that may not be as easily evaluated by a clinician. The modest ICC values found for the questionnaire components on non-PFPS subjects suggest that small fluctuations in these variables should normally be expected. Occasional knee complaints can occur in the normal population and the subjective nature of function and pain assessment relies on the individual's interpretation. Therefore, although important to consider, and certainly a limitation to the degree of validity, high reliability values using visual analogue scales for measurement of these variables are unlikely to be found. A compromise may be necessary in order to achieve a reasonable balance between the properties of reliability and responsiveness to clinical change.

Only five clinician rated tests were included in this study and based on the results of the internal consistency analyses, there is evidence that the tests are representative of one area of evaluation. The tests chosen were based on the results of a previous content validation process (Harrison, 1994). In the content validation study, 21 tests were rated by expert reviewers, considering the ability of each test to detect clinical change in PFPS patients. In determining the responsiveness of a test, one must establish whether the test appropriately discriminates between the diseased or injured individual and the healthy individual. Based on the results of the ANOVA, all clinical tests, except patellar mobility, differentiated between the populations. The scale used to describe patellar mobility ranged from hypomobility at one end, to hypermobility at the other end, with a normal value in the middle of the scale. In comparison, all other scales indicated normal at one end of the scale and abnormal at the other end of the scale. As ANOVA is dependent on the location of the mean (Cohen, 1988), the

insignificant results found in measurement of patellar mobility may be explained, in part, due to the nature of the scaling. It may be appropriate to redefine the scale, similar to the others, if group means are to be compared. It is interesting to note that patellar mobility was found to demonstrate reasonable properties of specificity when dichotomous classification was used. This further supports the need to refine the method of measuring patellar mobility.

When the clinical tests were investigated for sensitivity and specificity, all tests demonstrated high specificity but low to fair sensitivity. These findings suggest that the tests, when classified dichotomously, accurately discriminate between PFPS subjects and non-PFPS subjects. However, within the PFPS group, many of the tests may not be positive. Considering the lack of a unique etiology for the condition, these results are not surprising. In this study, at least one positive clinical test was found in all but one patient. The probability of a patient having more than one positive test was high, however, no one test seemed to be more strongly linked with PFPS. In regards to the higher specificity values found in this study, it must be remembered that clinicians were not blinded to subject groupings, therefore knowledge of the condition may have affected clinician rating. The standardization of protocols and operational definitions attempted to reduce the error that may have been present due to knowledge of the grouping. The reality, however, is that clinicians judge clinical findings in conjunction with the patient's history and subjective reporting of symptoms.

Of the various activity measurement methods evaluated in the study, there were limited relationships, especially when considering the combined activity pattern scores and the combined visual analogue questions with the LRCQ. The findings are not surprising when one considers the many factors that affect a person's activity level in a typical day. The demands of school, work, or family most likely have a greater influence on a person's actual activity level than a condition like PFPS. Many patients report that the problem is annoying and that they continue their activities of daily living despite the knee pain.

Although the three questions using the visual analogue scales demonstrated

modest internal consistency values, the best methods of activity rating that were found to discriminate between the PFPS and non-PFPS groups were: the question rating overall frequency of participation in intense physical activity with a visual analogue scale, and the four questions documenting the amount of time spent daily in standing, walking, squatting/kneeling and sports/recreation activities using ordinal scales. Explanations as to why only four of the eight activities demonstrated discriminative properties between groups are most likely related to the inability for many people to control all aspects of their activity level based only on pain or functional limitations. Sports and recreation activities are more commonly voluntary endeavours, rather than imposed by the job or school. As a result, these activities may be better discriminative indicators between the patient and non-patient populations. However, it must be emphasized that other factors, especially age of the subject and the time of year (season) can affect the level of activity. In this study, age was controlled by matching subjects from each group, and time of year was controlled, to some extent, by holding the study over a six month period and following subjects only over a one month period.

One may have thought that there should be a difference with stair climbing, as many patients complain of problems or pain in ascending or descending stairs due to the increased joint reaction force with this activity (Fulkerson & Hungerford, 1990). Stair activity is very much related to the environment in which one works or lives, and may be the reason that the groups did not differ in this activity. As a large number of both samples were students, the lack of difference in sitting is not surprising, as students have no choice but to sit during the school day. The lack of difference between time spent lying can be explained by the fact that this activity is primarily a measure of the amount of time sleeping. Perhaps with a condition where the level of disability was greater, or pain at rest was a common symptom, time spent lying may be affected. To summarize the role of activity measurement in PFPS subjects, it would appear that in this study, levels of participation in sports/recreation or intense physical activities may be useful discriminative measures of activity between PFPS populations and non-PFPS

populations. For those individuals who are not physically active, the activity pattern questions recording time spent standing, walking and squatting/kneeling may be better indicators, as these activities seem to better discriminate between the two groeps.

One question that this study attempted to answer was whether the three domains of evaluating PFPS, namely pain, function, and activity, were separate entities within the condition. The resu|ts of the correlation analysis using the MTMM and principal component analysis confirmed that pain and function represented only one domain while activity appeared to be unique from the other two. As noted by the name, Patellofemoral Pain Syndrome, pain is a primary component of the condition. Patients infrequently report night pain and as is well recognized, pain often increases with knee movement (Fulkerson & Hungerford, 1990). The significance of these findings is that for purposes of classification and evaluation of patients, function and pain scales can be combined without losing a unique dimension of the condition. The high internal consistency fount for the combined function and pain questions also supports this finding.

It also appears that the patient questionnaire is a better method for evaluating both activity and pain/function than the clinician method used in this study, as the clinician method appears to include method variance. The shared method variance may be explained because the evaluation takes place in a controlled clinical environment with patient signs and symptoms being observed over a brief period of time. Also, if a clinician observes a person walking up stairs, they also gather information about pain by watching the person's expressions, noting a hesitancy to weight bear on that side due to pain, and perhaps inability to do the activity because it hurts. The reason for shared method variance between clinician rated activity and pain and function may be due in part to the fact that the clinician rated activity level is based on the history of activity rather than direct observation. Although clinician were to rate activity according to reported frequency and intensity of participation in activities, what they observed during the performance of the pain and function testing may have influenced their

rating of activity level. These findings suggest that the questionnaire items were better able to discriminate between activity levels and pain/function measures than the clinician. Therefore, it would appear that if activity is to be considered as a unique entity in PFPS, the questionnaire method would be the preferred method of evaluation.

Determining guidelines for clinically relevant change is essential for evaluative research. Based on the results of this study, and in consideration of a previous study that investigated effect sizes for pain and function measurement on PFPS patients (Harrison, 1994), the following estimates of clinically relevant change have been recommended. For mean scores of the patient questionnaire rating of pain and function, an increment of 1 centimetre on a 10 centimetre scale appears to be a reasonable estimate of a clinically relevant change. The results of this study would suggest that activity measurement should be used to categorize subjects but not to monitor clinical change. As mentioned previously, activity levels are dependent on many factors unrelated to an individual's physical problems resulting from disease or injury. Considering the small effect sizes found for the two activity measures felt to be better discriminative indicators between the PFPS and non-PFPS populations, activity evaluation may be useful as an independent variable in evaluative studies, but not as a dependent variable. Interpretation of effect sizes for the five clinical tests are limited for two major reasons. First, there is not a criterion on which to measure improvement or deterioration. Secondly, not all tests are positive in patients, which limits the ability to determine clinically relevant change, as clinical tests may not change because they are not positive in the first place. Assuming that pain and function ratings by patients are appropriate criteria for change in the PFPS subjects, and based on the effect sizes found in the non-PFPS and PFPS groups, the following recommendations regarding the five clinical tests are made: 1) A one centimetre increment on a 10 centimetre visual analogue scale would be an estimate of clinically relevant change for lower extremity alignment, while 1.5 centimetre increments would be required to detect changes in flexibility of rectus femoris and tensor fascia latae

muscle groups; and 2) Due to the medium effect sizes found for patellar orientation and patellar mobility in the non-PFPS group, large effect sizes, equivalent to a 1.5 cm increment may be more appropriate estimates of clinically relevant change. Having proposed these guidelines, it must be pointed out that patellar mobility, as evaluated in this study, should be cautiously interpreted. Refinement in scaling, by including 10 cm. increments for both hypermobility and hypomobility, along with definitions for these scales may improve the discriminative properties of the test. The finding of a medium effect size for patellar orientation for both PFPS and non-PFPS groups, suggest that refinement in methods of measurement will be needed prior to accepting this clinical test as a useful measure. This clinical test is plagued with tester bias, and methods that have been developed to improve the objectivity of measuring patellar alignment (Ehrat, Edwards, Hastings & Worrell, 1994) have demonstrated very limited test-retest reliability.

A final issue that needs to be addressed is the use of the clinical tests and the questionnaire items as a cumulative index. Intuitively one may assume that a cumulative score may decrease the sensitivity of a measure and limit the information that may be gathered regarding clinical change. In this study combined scores of the clinical tests, excluding patellar mobility, and patient ratings on pain and function demonstrated high levels of responsiveness. Activity was also excluded from the cumulative index due to the limitations previously discussed. A cumulative index may be useful when investigating effects of treatment on groups of patients. In a clinical environment this method of recording would also be very practical.

Conclusions

Pain and function appear to represent one area of PFPS evaluation, with self-reporting methods, as described in this study, found to reasonably discriminate between patient and non-patient populations. For evaluative purposes, the questionnaire items can be used individually or combined, depending on the nature of the research, as they demonstrate the ability to detect

clinically relevant change. Specific questionnaire methods of evaluating activity levels in PFPS subjects were found to be primarily useful in categorizing patients, with limited ability to detect clinical change. Modest test-retest reliability for function, pain and activity measurement using self-reporting methods must be remembered when establishing clinical guidelines for changes in these variables.

Of the five clinical tests analysed, lower extremity alignment and flexibility of rectus femoris and tensor fasciae latae muscle groups demonstrated reasonable discriminative diagnostic properties, in addition to responsiveness. Refinement of measurement methods for patellar mobility and better knowledge of the role of patellar mobility in PFPS are required before this variable should be accepted as an appropriate outcome measure. Patellar orientation, although able to discriminate between patient and non-patient populations, demonstrated limited stability. Poor test-retest reliability may be responsible for the inability for the test to detect true clinical change to be appropriately detected.

A cumulative index representing clinical tests, excluding patellar mobility, and patient ratings of pain and function may be useful outcome measures in the clinical environment. In addition, both clinical tests and patient self-reports provide useful information on which to develop treatment programs. It must be emphasized that lack of a unique etiology and presentation for this condition makes it impossible to develop one tool that will be appropriate for all individuals. Therefore, selection of the patients and clear identification of the purpose of evaluation must always be considered prior to determining the appropriate measurement method.

In summary, with PFPS patients presenting with abnormal lower extremity alignment and inflexibility of the hip flexor and abductor muscle groups, the clinical tests uset in this study appear to be appropriate and sensitive measures. Patellar mobility, as defined in this study, was not an appropriate measure for the patient population. Further work is required to define and indeed investigate whether patellar mobility is an important indicator for PFPS. With development, to improve consistency of measurement, patellar orientation has a role to play in evaluation.

The patient questionnaire used in this study is a useful addition to clinical tests as it objectively meacures the subjective component. This study would suggest that the questionnaire does not differentiate pain symptoms from functional limitations. In the PFPS population the two variables appear to represent the same phenomena. If the length of the questionnaire is of concern, this finding provides rationale for the clinician to ask the patient to rate pain or function only, on the visual analogue scales, without missing an important component of the condition. An exception to this may exist in populations with extreme levels of activity. For example, these findings may not be generalized to the elite level athlete or the very sedentary individual. In the future the questionnaire must be investigated on these populations.

In large population studies, a cumulative index using all components of the evaluation would be satisfactory, if the groups were similar in regards to the clinical presentation. In the clinical setting, however, the clinician may find that the tests and questionnaire provide more information if they are analysed separately. Information regarding specific clinical findings or the patient's subjective report are the basis on which the clinician monitors the individual and determines appropriate treatment.

TABLE III-1. Inclusion and exclusion criteria for PFPS subjects.

INCLUSION:
- 12 to 45 years of age
- ambulation without external support
- complaint of patellofemoral pain with activity
- patellar pain on palpation and with resisted knee extension

EXCLUSION:
- positive neurological signs/symptoms
- recent injury to other lower extremity joints or spine/pelvis
- knee injury including ligamentous or meniscal involvement:
  surgically treated or
  restricting activity for more than 3 months
- extensor apparatus involvement (quadriceps or patellar tendinitis)
- bursitis or isolated inferior fat pad involvement
- reflex sympathetic dystrophy
- radiological changes of the knee associated with:
  osteoarthritis
  osteochondral or chondral fractures
- recent steroid injections to the knee (1 year)
- gross knee effusion

TABLE III–2. Demographic data for PFPS and Non–PFPS groups

Group	N	Age (yrs) $\bar{X}$, (SD) (range)	Ht (cm) $\bar{X}$, (SD) (range)	Wt (kg) $\bar{X}$, (SD) (range)	Gender (No) Female	Male
PFPS 1	41	24.1 (8.4) (12–48)	169.8 (9.0) (154–188)	70.7 (15.13) (48–125)	26	15
PFPS 2	28	19.7 (4.3) (12–26)	169.8 (7.9) (154–188)	68.7 (16.23) (48–125)	18	10
Non–PFPS	28	19.0 (4.0) (13–27)	167.2 (5.1) (159–180)	61.2 (9.9) (48–88)	24	4

PFPS 1 – All patients; PFPS 2 – Patients matched with Non–PFPS subjects by age

TABLE III-3. Clinician rating of eight variables using visual analogue scales

Test	PFPS group (n=28) $\bar{X}$ (SD) (range) Affected Leg	PFPS group (n=28) $\bar{X}$ (SD) (range) Non-Affected Leg	Non-PFPS group (n=28) $\bar{X}$ (SD) (range) Right Leg	Non-PFPS group (n=28) $\bar{X}$ (SD) (range) Left Leg
Lower Extremity Alignment	3.83(1.75)(0-7)	3.03(1.94)(0-7)	1.21(1.5)0-6)	1.46(1.5)(0-6)
Patellar Mobility	5.14(2.2)(1-9)	5.0(2.09)(1-9)	4.82(0.19)(0-7)	4.96(0.64)(3-6)
Patellar Orientation	3.21(2.01)(0-7)	2.52(1.66)(0-6)	0.18(0.39)(0-1)	0.25(0.52)(0-2)
Flexibility Rectus Femoris	3.34(3.24)(0-9)	2.79(2.77)(0-8)	0.32(1.19)(0-5)	0.39(1.23)(0-5)
Flexibility Tensor Fasciae Latae	4.69(2.62)(0-9)	3.93(2.4)(0-8)	0.5(1.04)(0-4)	0.5(1.04)(0-4)
Activity Level	3.85(1.86)(0-8)		0.75(1.35)(0-4)	
Pain	3.22(1.02)(1.17-5.33)		0.04(0.19)(0-1)	
Function	3.74(2.25)(0-8.45)		0.04(0.19)(0-1)	

Visual analogue scale: 0 to 10 cm.: 0 - normal, 10 - severely affected; except patellar mobility: 0 - hypomobile, 5 - normal, 10 - hypermobile; and activity level: 0 - very high to 10 - very low

TABLE III-4A. Visual analogue scores for questionnaire components for matched PFPS and Non-PFPS groups

Visual Analogue Scores	PFPS Group (n=28)	Non-PFPS Group (n=28)
	$\bar{X}$ (SD) (range)	$\bar{X}$ (SD) (range)
1 Six Pain Questions*	3.63 (0.95) (2-5.33)	0.54 (0.57) (0-2.17)
2 Three Activity Questions:	3.72 (2.12) (0-8.45)	0.60 (0.79) (0-2.33)
3 Thirteen Function Questions*	3.72 (1.92) (0-8)	0.46 (0.77) (0-3.62)

Visual analogue scale 0 to 10 cm.:

1 0 - no pain, 10 - severe pain

2 0 - very high, 10 - very low

3 0 - no problem, 10 - unable to do

* sig. $p < .05$ (ANOVA)

TABLE III-4B. Activity ratings for Lipid Research Clinic Questionnaire (LRCQ) and Activity Pattern Questions for matched PFPS and Non-PFPS groups

Activity Ratings	PFPS (n=28)	Non-PFPS (n=28)
	$\bar{X}$ (SD) (range)	$\bar{X}$ (SD) (range)
LRCQ Score	1.93 (0.96) (1-4)	2.61 (0.5) (3-4)
Eight Activity Pattern Questions	24.96 (4.62) (15-35)	26.39 (2.39) (22-31)

LRCQ: 1 - very low active; 2 - low active; 3 moderately active; 4 - very active;
Activity Pattern: 0 - very low, 40 - very high)

TABLE III-5A. Internal consistency of five clinical tests and questionnaire components for pain, function and activity using Cronbach's Alpha Model for 41 PFPS subjects

Test Components	Corrected Item-total Correlation
Five Clinical Tests	0.66
Six Pain Questions	0.48
13 Function Questions	0.92
Three Activity Questions	0.22
Combined Pain/Function Questions	0.90

TABLE III-5B. Intra-class correlation coefficients (ICC) for five clinical tests and questionnaire components of pain, function and activity on 28 Non-PFPS subjects

Test Components	ICC
Five Clinical Tests	0.97
Six Pain Questions	0.71
13 Function Questions	0.60
Three Activity Questions	0.76

TABLE III–6A. Correlation matrix for clinician ratings (method 1) and subject ratings (method 2) of pain, function and activity in PFPS group (n=41)

		Clinician Rating			Subject Rating		
		Activity	Pain	Function	Activity	Pain	Function
Clinician Rating	Activity	1.00					
	Pain	0.43*	1.00				
	Function	0.43*	0.71*	1.00			
Subject Rating	Activity	0.49*	0.23	0.01	1.00		
	Pain	0.05	0.41*	0.61*	0.02	1.00	
	Function	0.05	0.42*	0.59*	0.05	0.44*	1.00
		* Sig < .01					

TABLE III–6B. Principal components factor analysis with varimax rotation of clinician ratings and subject ratings for pain, function and activity.

Factors	Eigenvalues	Cumulative Variance
Factor 1	2.89	48.2%
Factor 2	1.42	71.9%
Factor 3	0.60	82.0%
Factor 4	0.52	90.7%
Factor 5	0.37	96.9%
Factor 6	0.18	100%

Factor Loading Matrix

				Rotated Matrix	
		Factor One	Factor Two	Factor One	Factor Two
Clinician Rating	Activity	0.43	**0.79**	0.40	**0.81**
	Pain	**0.81**	0.24	**0.79**	0.28
	Function	**0.93**	0.07	**0.92**	0.11
Subject Rating	Activity	–0.35	**0.77**	–0.39	**0.76**
	Pain	**0.74**	–0.13	**0.74**	–0.10
	Function	**0.73**	–0.33	**0.74**	–0.29

TABLE III-7. Diagnostic discrimination of musculo-skeletal lower extremity tests for matched PFPS and Non-PFPS groups

Tests (n=56)	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value
Lower Extremity Alignment	57%	92%	88%	69%
Patellar Mobility	78%	86%	85%	80%
Patellar Orientation	39%	86%	73%	59%
Flexibility Rectus Femoris	50%	92%	88%	63%
Flexibility Tensor Fasciae Latae	67%	92%	90%	74%

TABLE III-8. Effect sizes for clinician rated tests and questionnaire components for PFPS and Non-PFPS groups

Tests	PFPS (n=28)	Non PFPS (n=28)
Clinician Ratings	Effect Size	Effect Size
Lower Extremity Alignment	0.53	0.18
Patellar Mobility	0.11	0.56
Patellar Orientation	0.37	0.4
Flexibility Rectus Femoris	0.73	0.28
Flexibility Tensor Fasciae Latae	1.22	0.22
Activity	0.1	0.1
Pain	0.8	0.44
Function	0.91	0
Questionnaire Scores		
Activity1 VAS	0.08	0.08
Activity2 VAS	0.58	-0.31
Activity3	0.13	0.08
Pain VAS (Six questions)	0.94	0.1
Function VAS (Thirteen questions)	0.54	0.34

Activity1 - One question - frequency of participation in sports/intense activity
Activity2 - Three activity questions
Activity3 - Four activity pattern questions (ordinal scale)

$$\text{Effect size} = \frac{\text{Final Test - Initial Test}}{\text{Pooled S.D.}}$$

TABLE III-9. Cumulative scores and effect sizes for two test periods for PFPS and Non-PFPS groups

	PFPS (n=28)		Non-PFPS (n=28)	
Tests	$\bar{X}$ (SD) (range)	Effect Size	$\bar{X}$ (SD) (range)	Effect Size
[1] Clinician Ratings				
Initial Test	14.52 (6.32) (3 - 25)		6.86 (3.10) (0 - 15)	
Final Test	8.86 (5.49) (2 - 20)	0.92	6.64 (1.97) (5 - 12)	0.20
[2] Subject Questions on Pain and Function				
Initial Test	69.45 (29.65) (19-125)		9.22 (12.38) (0 - 56)	
Final Test	47.55 (43.07) (0-131)	0.78	6.25 (8.93) (0 - 34)	0.28
[1,2] Clinician Ratings and Subject Questions on Pain and Function				
Initial Test	82.18 (33.79) (33-140)		16.07 (13.84) (3 - 68)	
Final Test	55.97 (39.73) (3 - 136)	1.0	12.89 (10.19) (5 - 44)	0.30

[1] Lower extremity alignment, patellar orientation, flexibility of rectus femoris and tensor fasciae latae muscle groups (Cumulative Score 0 to 40 cm)

[2] Six pain questions, Thirteen function questions (Cumulative score 0 to 190 cm)

BIBLIOGRAPHY

Ainsworth BE, Jacobs DR, Leon AS: Validity and reliability of self-reported physical activity status: The Lipid Research Clinics questionnaire. Med Sci Sports Ex 1993; 125(No.1): 92-98.

Campbell DT, Fiske DW: Convergent and discriminant validation by the multitrait multimethod matrix. PsycholBull 1959; 56: 81-105.

Chesworth BM, Culham EG, Tata GE, Peat M: Validation of outcome measures in patients with patellofemoral syndrome. JOSPT 1989; 11: 302-308.

Cohen J: Statistical Power Analysis for the Social Sciences. 2nd ed. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1988.

Donatelli R: The Biomechanics of the Foot and Ankle. F.A.Davis, 1990.

Dunteman GH: Principal component analysis. In: Series on Quantitative Applications in the Scoia Sciences. Sage Publications, 1989; Beverly Hills: 7-13.

Ehrat M, Edward J, Hastings D, Worrell T: Reliability of assessing patellar alignment: The A Angle. JOSPT 1994; 19(1): 22-27.

Fairbank JCT, Pynsent PB, van Poortvliet JA, Phillips H: Mechanical factors in the incidence of knee pain in adolescents and young adults. J Bone Joint Surg [Br] 1984; 66B: 685-693.

Ferketich SL, Figueredo AJ, Knapp TR:Focus on psychometrics:The Multitrait-Multimethod Approach to construct validity. ResNurs& Health 1991; 14: 315-320.

Figuerdo AJ, Ferketich SL, Knapp TR: Focus on psychometric more on MTMM: The role of confirmatory factor analysis. ResNurs& Health 1991; 14:387-391.

Fletcher RH, Fletcher SW, Wagner EH: Chapter 3 Diagnosis In: Clinical Epidemiology. 2nd Ed. Baltimore: Williams & Wilkins, 1988: 42-75.

Fulkerson JP, Hungerford DS: Disorders of the Patellofemoral Joint. 2nd ed. Baltimore: Williams and Wilkins, 1990.

Goldberg B: Chronic anterior knee pain in the adolescent. PediatrAnn 1991; 20: 186-7,190-3.

Harrison EL: Development of a clinical tool and patient questionnaire for

evaluation of patellofemoral pain syndrome patients. Unpublished manuscript 1994.

Hefzy MS, Jackson WT, Saddemi SR, Hsieh Y-F: Effects of tibial rotations on patellar tracking and patello-femoral contact areas. JBiomedEng 1992; 14: 329-343.

Jette AM: Measuring subjective clinical outcomes. Phys Ther 1989; 69: 580-584.

Kaplan RM, Bush JW, Berry CC: Health status: types of validity and the index of well-being. Health Serv Res 1976; 11: 478-507.

Kendall FR & Kendall McCreary E: Muscles Testing and Function. 3rd ed. Baltimore :Williams & Wilkins Co., 1983: 160-168.

Magee DJ: Orthopedic Physical Assessment. 2nd ed. Philadelphia: W.B.Saunders Co., 1992.

McConnell J: The management of chondromalacia patellae: A long term solution. AustralJPhysiother 1986; 32: 215-223.

Nordin M, Frankel BH: Basic Biomechanics of the Musculoskeletal System. 2nd ed. Philadelphia: Lea & Febiger, 1989.

Noyes FR: The Noyes Knee Rating System An assessment of subjective, objective, ligamentous, and functional parameters. Internal Publication of Cincinnati Sportsmedicine Research and Education Foundation. Cincinnati, Ohio (1990).

Payton OD: Research: The Validation of Clinical Practice. 2nd ed. Philadelphia: F.A.Davis Co., 1988.

Reid DC: Sports Injury Assessment and Rehabilitation. New York: Churchill Livingstone, 1992.

Travers RMW: Rational hypotheses in the construction of tests. Educational and Psychological Measurement 1951; 11: 128-137.

CHAPTER 5
DISCUSSION AND CONCLUSIONS

The development of a clinical evaluation method involves various stages. The three studies investigated a number of psychometric properties for PFPS evaluation methods, including certain types of validity. It must be emphasized that evidence for validity is developed on a continuum. In clinical research, the status of health measurement systems and outcome measures is constantly changing. Expanding knowledge in areas such as anatomy, physiology, in addition to disease and injury pathology, etiology, clinical presentation and treatment, contribute to the demand to continually evaluate and modify existing tools and develop new tools.

The lack of unique etiology or pathology in PFPS challenges the clinician to develop evaluation methods that can accommodate to the individual requirements of each patient. Attention to clinimetrics (Feinstein, 1987), which is the quantification of clinical data such as that observed and reported by clinicians, is essential in PFPS, as the evaluation is primarily centred on clinical phenomena. In addition, patient self-reports which have been transformed into quantifiable data must be considered as important components of the evaluation.

Three studies were performed over a four year period. The first study investigated properties of reliability and validity for existing clinical evaluation methods. In the second study, a content validation process evaluated the components of the tool by analysing expert clinicians' opinions. Based on the first two studies, a PFPS measurement tool including five clinical tests and a patient questionnaire was developed. The tool was used to assess a group of PFPS subjects and a group of individuals without the condition in the third study. Evidence for validity of the different components of the tool considered the results of several analyses including: test-retest reliability, internal consistency, known group differences, correlational analysis, principal components analysis, sensitivity, specificity and effect sizes. Individually these properties can provide information that will assist in determining construct validity of a tool. The best construct is

defined as the one around which one can build the greatest inferences in the most direct fashion (Cronbach & Meehl, 1955).

Consistency

Test-retest

The results of these studies suggest that test-retest reliability values for the pain and function questionnaires using visual analogue scales are modest. It must be emphasized that these results were found in both patient and healthy subjects. Even with changes from study one to study three in the number and wording of the pain and function questions, and in the scaling of the function questions, reliability coefficients were similar. Small fluctuations in pain and function on a 10 cm. visual analogue scale, although clinically insignificant, represented decreased consistency using statistical methods such as intra-class correlation coefficients. Compromises are often necessary in order to balance the essential properties of an evaluation method. For example, a tool that is highly reliable may not be sensitive to clinical change (Guyatt, Walter & Norman, 1987). The results of the clinical studies suggest that the consistency of the pain and function questionnaire items is reasonable for the purposes of evaluative studies. However, clinically significant change must be determined prior to interpreting results of clinical trials.

There are many factors that can influence pain and function on a day to day basis including cultural and individual factors (Melzack, 1980). To eliminate these variables would be difficult and certainly impractical. Function is influenced by occupation, environment and lifestyle. As a result of these factors, it seems reasonable that only modest consistency will be found when evaluating pain and function. Improved reliability may be found if a more controlled environment is used to evaluate components of function, for example walking on a treadmill. The problem with this approach is that function implies the ability to carry out activities in ones normal environment. The clinical environment limits the validity of measurement by jeopardizing the ability to measure the true phenomena. Clinicians must consider the small fluctuations that normally occur in PFPS

patients in a stable situation, when determining clinically relevant changes for pain and function. The self-reports of pain and function, as described in these studies, can be effective evaluative measurements. The practicality and the ability to measure the true phenomena, not only consistency, must be kept in mind when developing evaluative methods.

The five clinician rated tests including lower extremity alignment, patellar orientation, flexibility of tensor fasciae latae and rectus femoris muscle groups and patellar mobility demonstrated high test-retest reliability in the non-patient population. These results are promising in regards to tester consistency in non-symptomatic subjects. Future work must investigate the inter-rater and intra-rater reliability of these clinical tests in PFPS populations with various levels of severity. For example, consistency in evaluating patients with severe abnormality of lower extremity alignment or patellar orientation may be different than consistency for evaluating minimal abnormality. Information on normal standards for each of the clinical tests is needed to improve the interpretation of findings.

Internal Consistency

The questionnaire items for function demonstrated high internal consistency in both clinical trials. The original functional index questionnaire (Chesworth, Culham & Tata, 1989) consisting of eight questions, using an ordinal scale with three responses, was revised for the second clinical trial. Five questions were added to investigate functional limitations associated with a more active, athletic lifestyle. Visual analogue scales were used to rate each of the 13 questions. The pain question items showed lower internal consistency, however, only 3 and 6 questions were rated in study 1 and study 3 respectively, which can account in part for the lower homogeneity. When pain and function items were combined, a high internal consistency was found which provides support for combining these areas in evaluation.

The five clinical tests: lower extremity alignment, patellar mobility, patellar orientation, and flexibility of rectus femoris and tensor fasciae latae muscle groups showed modest internal consistency. The findings of modest internal consistency

suggest that a relationship exists between the five clinical tests. In the first study, a previously developed patellofemoral scale (Reid, 1992) consisting of a combination of clinical tests and various functional components demonstrated similar properties of modest internal consistency. The condition, as commonly reported (Reid, 1992, Goldberg, 1991), does not present with one clinical presentation, but potentially a number of clinical signs and symptoms in different combinations. The clinical tests measure various aspects of the condition focusing on biomechanics of the lower extremity, patellofemoral joint articulation and muscle flexibility. The modest internal consistency of the five clinical tests in addition to the high test-retest reliability suggest that these evaluation methods have a role to play in evaluative research.

Known Group Differences

All components of the questionnaire, except for activity level, and all clinical tests, except for patellar mobility, were effective at differentiating between the PFPS population and the non-affected population. Known group difference is considered a minimal requirement for validation (Cronbach & Meehl, 1955). The reason that activity level did not differentiate between the two populations may be explained by the multiple factors that influence activity. People maintain activity level for reasons not directly controlled by a painful condition. Social, psychological, economic, cultural and environmental factors affect normal activity levels (Fordyce, 1983, Melzack, 1983). Patients often report that although the condition is annoying that they continue to participate in activities. Therefore gathering information on activity level and intensity of activity appears to be useful primarily for categorizing individuals. Function and pain measures demonstrate better discriminative properties.

The patellar mobility test did not distinguish between PFPS and non-PFPS populations. Although the expert reviewers rated patellar mobility as an important indicator for detecting clinical change, there does not appear to be strong evidence in the literature to directly link patellar mobility with PFPS. Changes in mobility of the patella are more commonly reported in patients with dislocation or

subluxation of the patella (Reid, 1992; Kolowich, Paulos, Rosenberg et al., 1990). Passive movements of the patella primarily evaluate the integrity of non-contractile elements surrounding the joint. Although these passive stabilizers play a role in normal biomechanics of the patellofemoral joint, the nature of PFPS suggests that dynamic control from muscle integrated with connective tissue, is perhaps more important. Further investigation is needed to determine the incidence of dysfunction of contractile and non-contractile components of the patellofemoral joint.

Correlation Analysis

Based on the content validation process it was theorized that three distinct domains of pain, function and activity would be found in the PFPS population using the questionnaire method. The MTMM (multi-trait multi-method) analysis and principal component analysis supported the concept that the questionnaire items primarily represented two domains: pain and function combined as one domain and activity the second domain. Unlike other conditions where signs like swelling may be present, PFPS patients' primary complaint is pain associated with and affected by certain functions. Therefore, when evaluating PFPS patients, pain and function questions can be successfully combined without losing a separate dimension of the condition. The finding that the activity questionnaire items represent a unique domain, unrelated to pain and function, adds further support to the finding that activity level may not be an appropriate evaluative measure. In developing studies to evaluate treatment efficacy in PFPS patients, these results would suggest that activity level should be considered as an independent variable rather than an outcome measure.

One additional property investigated by correlational analysis was the criterion validity of the activity questions. The findings suggest that the questions that best measured activity level in the PFPS population were frequency of participation in intense activities and the amount of time spent standing, walking, squatting/kneeling and participating in sports/recreation.

Sensitivity and Specificity

Although the majority of non-PFPS subjects did not present with positive findings in any of the five clinical tests, there was only a small ratio of patients that presented with positive findings for all tests. As a result the sensitivity of the clinical tests was only fair to modest while specificity was high. As mentioned previously, the lack of unique etiology is a primary reason for these findings. Clinicians must consider the findings for each patient and determine clinical change based on the initial positive tests. Although the cumulative index for clinical tests was found to be reasonable at detecting change, in small studies or single case analysis, combining tests will decrease the ability to detect change if only a small number of tests are positive.

Treatment programs are commonly directed by positive clinical tests. For example, a patient with tightness of the tensor fasciae latae muscle group will be prescribed a stretching program. In some cases, this may be the only positive finding other than the subjective reports of pain and functional limitations. It is important for the clinician to identify the key indicators, if possible, prior to treatment, in order to direct treatment appropriately. Also, it is essential that treatment programs are not discarded or promoted as a result of not measuring the appropriate outcomes.

The findings of these studies suggest that the clinician can appropriately use the function and pain questionnaire items for monitoring change in subjective variables in PFPS patients. The more detailed function questionnaire may be more appropriate for active individuals, while the original functional index questionnaire (FIQ) (Chesworth et al., 1989) appears to be a satisfactory measure for individuals who do not pursue recreationally active lifestyles.

In regards to clinical tests, lower extremity alignment and flexibility of tensor fasciae latae and rectus femoris muscle groups appear to be appropriate evaluative measures for the condition. Patellar orientation appears to have potential as an evaluative measure. The patellar measurements reported in the literature, namely the Q-angle (Ficat & Hungerford, 1977) and more recently the A-angle (Arno, 1990), present with large variations in the normal population. The

method for determining patellar orientation in this study was based on the relative position of the medial and lateral borders of the patella, the inferior and superior borders of the patella and the vertical position of the patella in relation to the femur. The three components in combination were considered when judging patellar abnormality. With some refinement of the definitions, this technique may be a better measurement method, as it considers more elements of the patellar position. With advances in imaging techniques, patellofemoral joint mechanics can be more closely evaluated considering the resting position of the patella. If the relationship between the resting position of the patella and patellar tracking can be more specifically defined, this will improve the clinical diagnosis. Imaging techniques may be useful in validating the clinical method of measuring patellar orientation. With the increased costs of diagnostic testing, the clinician must rely on simple tests that can be easily carried out in the clinical setting. Advanced imaging techniques will not be available or indicated in the majority of cases, thus valid clinical tests must be emphasized in training of health practitioners.

Passive patellar mobility, as measured in this study, did not distinguish between the PFPS and non-PFPS groups. There are many limitations to measurement of passive patellar mobility: lack of normal standards, differences in amount of manual pressure in the technique and the judgement of tissue tension, and the fact that both hypermobility and hypomobility are considered abnormal. It would appear from the findings in this study, that patellar mobility is not an important evaluative measure for PFPS.

For the purposes of this project, only five of 21 clinical tests were analysed. Expert reviewers were asked to rank the tests considering the ability of each test to detect clinical change in PFPS patients. Based on these results, future research may be directed towards determining the value of other tests such as gait, muscle tone of the quadriceps and flexibility of other muscle groups. Considering that lower extremity alignment was found to be a useful evaluative measure, gait assessment should be directed towards the components of normal lower limb biomechanics. In support of this approach, Chesworth et al. (1989)

suggested that gait analysis must be more comprehensive, as the basic elements such as stride length and cadence are not sensitive to clinical change. Evaluation of muscle tone of the quadriceps has been related to the dynamic control of patellar tracking (Souza & Gross, 1990). Clinical evaluation of muscle tone introduces the need for special instrumentation to measure the electrical activity of the muscle, therefore this may not be a practical test in the normal clinical setting. Flexibility of the hamstring and hip flexor muscle groups do not require special equipment and may be more easily implemented as useful tests, as basic test positions and normal standards have been established (Kendall & Kendall McCreary, 1983). Information on whether PFPS subjects generally have poor lower extremity flexibility, or if tightness of specific muscle groups is more evident, will assist in focusing treatment and provide the basis for prevention programs.

It must be stressed that although statistically significant differences were found between PFPS and non-PFPS groups for certain tests, the clinical relevance of these results must be carefully considered. The use of effect sizes is one method that can provide guidelines for determining important clinical change. Another aspect of clinical relevance is to consider developing outcome measures based on individual patient goals. The subjective tools and methods used in this study may be modified to reflect the patient's deficits, by considering what the patient reports as their limitations and goals.

In conclusion, the components of the PFPS evaluation method used in these studies appear to have merit in evaluative research and for monitoring clinical change of individual patients. As the patients in the studies were not representative of very severe cases of PFPS, the tool can not be generalized to all PFPS patients. Evidence has been provided to suggest that the questionnaire for pain and function, and the clinical tests for lower extremity alignment, flexibility of rectus femoris and tensor fasciae latae muscle groups are appropriate components of evaluation. Further research is required to refine the measurement of patellar orientation. In addition, other clinical tests should be evaluated to determine their relevance to PFPS.

BIBLIOGRAPHY

Arno S: The A angle: A quantitative measurement of patella alignment and realignment. JOSPT 1990; 12(6): 237-242.

Chesworth BM, Culham EG, Tata GE, Peat M: Validation of outcome measures in patients with patellofemoral syndrome. JOSPT 1989; 11: 302-308.

Feinstein AR: Clinimetrics. New Haven CT: Yale Univ Press, 1987.

Cronbach LJ , Meehl PE: Construct validity in psychological tests. Psychological Bulletin 1955; 52: 281-302.

Ficat RP, Hungerford DS: Disorders of the Patellofemoral Joint. Baltimore: Williams and Wilkins, 1977.

Fordyce WE: The Validity of Pain Behavior Measurement. In: Melzack R, ed. Pain Measurement and Assessment. New York: Raven Press, 1983:145-153.

Goldberg B: Chronic Anterior Knee Pain in the Adolescent. Pedatri Ann1991; 20: 186-7,190-3.

Guyatt G, Walter S, Norman G: Measuring change over time: assessing the usefulness of evaluative instruments. JChronDis 1987; 40: 171-178.

Kolowich PA, Paulos LE, Rosenberg TD, Farnsworth S: Lateral release of the patella: Indications and contraindications. Am J Sports Med 1990; 18:359-365.

Kendall FR, Kendall McCreary E: Muscles Testing and Function. 3rd ed. Baltimore: Williams & Wilkins, 1983.

Melzack R, ed: Pain Measurement and Assessment. New York: Raven Press, 1980.

Melzack R, ed: Pain Measurement and Assessment. New York: Raven Press, 1983.

Reid DC: Sports Injury Assessment and Rehabilitaiton. New York: Churchill Livingstone, 1992.

Souza DR, Gross MT: Comparison of vastus medialis obliquus: vastus lateralis muscle integrated electromyographic ratios between healthy subjects and patients with patellofemoral pain. Phys Ther 1991; 71: 310-316.

APPENDIX A-1 AND A-2

OCCUPATIONAL RATING SCALE BY WORK TITLE

(NOYES, 1990)

WORK ACTIVITY RATING AND LEVEL

(NOYES, 1990)

Occupational Rating Scale By Work Title

	0 = Sick leave/disability
Light	1 = Receptionist, data entry, telephone operator
Light	2 = Secretary
Moderately Light	3 = Studeint, file clerk
Moderately Light	4 = Salesman, inspector
Moderate	5 = Mail carrier
Moderate	6 = Police officer, stock person, nurse
Moderately Heavy	7 = Truck driver, fireman
Moderately Heavy	8 = Dock worker, plumber, carpenter
Heavy	9 = Farmer, sanitary garbage collectors
Heavy	10 = Construction, miner, logger

Work Activity Rating and Level

Work Activity Rating*	Work Activity Level
0	Disabled
1 - 20	Very light (VL)
21 - 40	Light (L)
41 - 60	Moderate (M)
61 - 80	Heavy (H)
81 - 100	Very Heavy (VH)

*Based on the total points scored for the types of activity performed at work

APPENDIX A-3

SPORT ACTIVITY AND FUNCTION FORM

(NOYES, 1990)

Patient Name | Involved Knee: Right ___ Left ___ | Date of Visit: mo ___ day ___ yr ___

Sports Activity Scale

Check the box which describes your level of sports activity before your original knee injury. Then, check the box which describes your level of sports activity at this time.

Original Activity		Current Activity
☐	100	☐
☐	85	☐
☐	80	☐
☐	85	☐
☐	80	☐
☐	75	☐
☐	85	☐
☐	80	☐
☐	75	☐
☐	40	☐
☐	20	☐
☐	0	☐

Level I (participates 4-7 days/week)
Jumping, hard pivoting, cutting (basketball, volleyball, football, gymnastics, soccer)
Running, twisting, turning (tennis, racquetball, handball, ice hockey, field hockey, skiing, wrestling)
No running, twisting, jumping (cycling, swimming)

Level II (participates 1-3 days/week)
Jumping, hard pivoting, cutting (basketball, volleyball, football, gymnastics, soccer)
Running, twisting, turning (tennis, racquetball, handball, ice hockey, field hockey, skiing, wrestling)
No running, twisting, jumping (cycling, swimming)

Level III (participates 1-3 times/month)
Jumping, hard pivoting, cutting (basketball, volleyball, football, gymnastics, soccer)
Running, twisting, turning (tennis, racquetball, handball, ice hockey, field hockey, skiing, wrestling)
No running, twisting, jumping (cycling, swimming)

Level IV (no sports)
I perform activities of daily living without problems
I have moderate problems with activities of daily living
I have severe problems with activities of daily living; on crutches, full disability

Highest Level (before injury) ___ / 100
Highest Level (current) ___ / 100

Change in Sports Activities

Check the box which best describes any change you have had in sports activities since your injury / surgery.
My sports activities have:

Not Changed
If yes, check one box below:
☐ I have no / slight problems (*c*)
☐ I have moderate / significant problems (*d*)

Decreased
If yes, check one box below:
☐ I now have no / slight problems (*e*)
☐ I now have moderate / significant problems (*d*)
☐ For reasons not related to my knee (*b*)

Stopped -- given up all sports
If yes, check one box below:
☐ I have moderate / significant problems when I play sports (*f*)
☐ For reasons not related to my knee (*g*)

Level _____

Function ADL

Check the problems you have during:

1. Walking
check one box:
100 ☐ normal, unlimited
95 ☐ some limitations
85 ☐ only 3-4 blocks possible
0 ☐ less than 1 block; cane, crutch

2. Stairs
check one box:
100 ☐ normal, unlimited
95 ☐ some limitations
85 ☐ only 11-30 steps possible
0 ☐ only 1-10 steps possible

3. Squatting / kneeling
check one box:
100 ☐ normal, unlimited
95 ☐ some limitations
85 ☐ only 6-10 possible
0 ☐ only 0-5 possible

Level _____ / 2x_____

Function Sports

Check the problems you have during:

1. Straight running
check one box:
100 ☐ fully competitive
95 ☐ some limitations, guarding
85 ☐ definite limitations, half speed
0 ☐ not able to do

2. Jumping / landing on affected leg
check one box:
100 ☐ fully competitive
95 ☐ some limitations, guarding
85 ☐ definite limitations, half speed
0 ☐ not able to do

3. Hard twists / cuts / pivots
check one box:
100 ☐ fully competitive
95 ☐ some limitations, guarding
85 ☐ definite limitations, half speed
0 ☐ not able to do

Level _____ / 2x_____

Problems with Sports

Describe the problems you would have with your knee after participating for one hour without guarding or limitations in each of the three sports categories below. (___check here if you are using a brace.)

Strenuous Sport
(soccer, football, basketball, volleyball)
check one box:
100 ☐ no problems
☐ moderate problems during or after game
☐ severe problems; cannot participate

Moderate Sport
(tennis, racquetball)
check one box:
95 ☐ no problems
☐ moderate problems during or after game
☐ severe problems; cannot participate

Light Sport
(golf, bowling, hiking)
check one box:
95 ☐ no problems
85 ☐ moderate problems during or after game
0 ☐ severe problems; cannot participate

Total Points _____

SPORTS ACTIVITY AND FUNCTION FORM — CINTI KNEE RATING SYSTEM

APPENDIX B-1

CINCINNATI KNEE RATING SYSTEM

(NOYES, 1990)

Patient Name		Injury Date	Wks Post-Op	☐ F ☐ M	Age	Occupation	Date
Date Surgery/ Procedure							
History							
Symptoms							

		RIGHT					LEFT			
		(4)	(3)	(2)	(0)		(4)	(3)	(2)	(0)	
General	1. Effusion	NL	Mild	Mod	Sev		NL	Mild	Mod	Sev	
	2. Total Flexion	NL	110°	90°	<90°		NL	110°	90°	<90°	
Points (16)	3. Lack of Extension	NL	5-10°	11-15°	>15°	____ R	NL	5-10°	11-15°	>15°	____ R
Rx__ Lx__	4. Quadriceps Weakness	NL	Mild	Mod	Sev	____ cm	NL	Mild	Mod	Sev	____ cm
Tibio Femoral	5. Joint Line Pain	NL	Mild	Mod	Sev	☐ MJL ☐ LJL	NL	Mild	Mod	Sev	☐ MJL ☐ LJL
	6. Crepitus	NL	Mild	Mod	Sev	____° range	NL	Mild	Mod	Sev	____° range
Points (12)	7. Compression Pain	NL	Mild	Mod	Sev	☐ MJL ☐ LJL	NL	Mild	Mod	Sev	☐ MJL ☐ LJL
Rx__ Lx__											
Patello Femoral Joint	8. Crepitus	NL	Mild	Mod	Sev	____° range	NL	Mild	Mod	Sev	____° range
	9. Compression Pain	NL	Mild	Mod	Sev		NL	Mild	Mod	Sev	
	10. Soft Tissue Pain	NL	Mild	Mod	Sev		NL	Mild	Mod	Sev	
Points (16)	Location						Location				
Rx__ Lx__	11. Soft Tissue Swelling	NL	Mild	Mod	Sev		NL	Mild	Mod	Sev	
	Location						Location				
Alignment	12. Lat. Sublux at 20° (% Width)	0-25	26-50	51-75	>75		0-25	26-50	51-75	>75	
	13. Med. Sublux at 20° (mm)	15	11-15	6-10	0-5		15	11-15	6-10	0-5	
Points (16)	14. Q Angle at 5°	0-15	16-20	21-25	>25		0-15	16-20	21-25	>25	
Rx__ Lx__	15. Q Angle Max at 20°	25	30	35	>35		25	30	35	>35	

Subluxation

Test	Right	Left	Difference	Test	Right	Left	Difference
Ant 25°	____mm	____mm	____mm	Med 0°	____mm	____mm	____mm
Ant 90°	____mm	____mm	____mm	Med 25°	____mm	____mm	____mm
P.S. (0-3)	____	____	____	Lat 0°	____mm	____mm	____mm
Post 25°	____mm	____mm	____mm	Lat 25°	____mm	____mm	____mm
Post 90°	____mm	____mm	____mm	ER 25°	____deg	____deg	____deg
RPS (0-3)	____	____	____	ER 90°	____deg	____deg	____deg

X-ray

	NL	Mild	Mod Narrowing <1/2 Joint	Sev Narrowing >1/2 Joint		NL	Mild	Mod Narrowing <1/2 Joint	Sev Narrowing >1/2 Joint	
16. Mid Tibiofemoral	☐	☐	☐	☐	☐ Sublux	☐	☐	☐	☐	☐ Sublux
17. Lat Tibiofemoral	☐	☐	☐	☐	☐ Tilt	☐	☐	☐	☐	☐ Tilt
18. Patellofemoral	☐	☐	☐	☐		☐	☐	☐	☐	
19. Alignment	____°			____ WBL		____°			____ WBL	

Points (12) Rx__ Lx__

Diagnosis	1.
	2.
	3.
Treatment	
Rehab	
Return to Activity	Impairment ☐ Full ☐ Partial ☐ Onset _____ ☐ End _____ ☐ Light Duty ☐ Full Duty

Date of Form	Next Appointment	M.D.	Rehab	Signature

KNEE EXAMINATION — CINN KNEE RATING SYSTEM

APPENDIX B-2

PATELLOFEMORAL FUNCTION SCALE

(REID, 1992)

Patellofemoral Function Scale

History

Pain	
None	10
During vigorous activity	8
During light activity	4
At rest after activity	2
Daily pain irrespective of activity*	0
Function	
Movie sign	
Absent	4
Present	0
Walking	
No restriction	5
Restricted	0
Stair climbing	
No restriction	5
Restricted	0
Jogging	
No restriction	6
Restricted	0
Sprint & cutting	
No restriction	10
Restricted	0
Orthosis**	
None	4
Knee sleeve or shoe insole	2
Total contact knee orthosis	1
Walking cane	0

Examination

Effusion	
None	6
Present	0
Patellar crepitus (passive)	
None	6
Mild	4
Severe***	0
Patellar crepitus (resisted)	
None	4
Mild	2
Severe	0
Quads atrophy	
None	4
>1cm	2
>2cm	1
>3cm	0
Apprehension Test	
Negative	6
Present	0
Clark's sign	
In extension	
Mild or Negative	4
Severe	0
At 10° flexion	
Mild or Negative	6
Severe	0
Range of motion	
Full	6
<10° limitation	4
>10° limitation	0
Pain with motion	
Full and pain free	6
Mild pain with resistance***	4
Severe pain with resistance	2
Painful no resistance	0
Squatting	
No problem	8
Slightly impaired	6
Not past 90°	2
Unable	0

80+	**Excellent**
70-80	**Good**
50-70	**Fair**
30-50	**Poor**
30	**Severe disability**

Needs stable knee
Only mark one in each category

*Includes night pain
**Regular use for activity
***Firm manual resistance
Compare to other knee

APPENDIX B-3

PATELLFEMORAL RATING SCALE

(SCHWARTZ, 1990)

PATELLOFEMORAL RATING SCALE

Name _____

Date _____

PAIN

none	10
during activity	7
after activity	3
at rest	0

FUNCTION

Cutting	no restriction	1
	restricted	0
Running	no restriction	1
	restricted	0
Walking	no restriction	1
	restricted	0
Stair climbing	no restriction	1
	restricted	0

(1) Total pain & function _____

SUBTRACTIONS

ROM-- subtract 1 point for each 10° loss compared to opposite knee _____

Cane or crutch-- subtract 1 point _____

EXAMINATION

Effusion	none	3
	present	0
Patellar crepitus	none	3
	minimal	2
	moderate	1
	severe	0

Quad atrophy (sitting, hip & knee flexed at 90°)

same as other thigh	3
convex but not equal	2
flat	1
concave	0

Quad tone	no loss	3
	minimal loss	2
	moderate loss	1
	severe loss	0

Tenderness on patellar compression

none	3
minimal	2
moderate	1
exquisite	0

(2) Total examination _____

TOTAL (1) + (2) _____

TOTAL SUBTRACTIONS _____

TOTAL P-F SCORE _____

APPENDIX C

FUNCTIONAL INDEX QUESTIONNAIRE

(CHESWORTH, CULHAM, TATA & PEAT, 1989)

NAME:

DATES TO RECORD:

FUNCTIONAL INDEX QUESTIONNAIRE

The following information is to be recorded at approximately the same time each day (preferably at bedtime).Put a check mark in the column that best describes the way you feel. Complete one box, with eight questions, for each date indicated at the top of this form. Fill in the date at the top of each box.

The following is an example of how to fill in the questions.

Today did you have any problem or discomfort in your ___ knee at all with the following activities?

	Unable to Do	Can do with Problem	No Problem	Unknown
Walking as far as a mile	( )	( )	( )	( )

Please complete the following:

Date:

Today did you have any problem or discomfort in your ___ knee at all with the following activities?

	Unable to Do	Can do with Problem	No Problem	Unknown
1. walking as far as a mile	( )	( )	( )	( )
2. climbing up 2 flights of stairs (16 steps)	( )	( )	( )	( )
3. squatting	( )	( )	( )	( )
4. kneeling	( )	( )	( )	( )
5. sitting for prolonged periods with your knees bent in one position	( )	( )	( )	( )
6. climbing up 4 flights of stairs (32 steps)	( )	( )	( )	( )
7. running a short distance, say 100 meters (about the length of a football field)	( )	( )	( )	( )
8. walking a short distance (about a city block)	( )	( )	( )	( )

APPENDIX D

CONTENT VALIDATION QUESTIONNAIRE

FOR EXPERT REVIEWERS

PATELLOFEMORAL CONTENT VALIDATION

QUESTIONAIRRE FOR EXPERT REVIEWERS

NAME:

ADDRESS:

PHONE:

FAX:

YEARS OF CLINICAL PRACTICE:

Please Circle the appropriate response:

I am interested in participating as an expert reviewer. YES NO

I will be available to respond to correspondence during the months of January and February. YES NO

How many new patellofemoral pain syndrome patients would you see on an average in your clinical practice each month?

Less than 4 4 to 8 8 to 12 more than 12

Please comment on any special interest, training, research or publications that you may have in the area of patellofemoral pain syndrome.

Thankyou! Please return in envelope enclosed.

January 27, 1993

Dear Expert Reviewer:

Thank you for agreeing to participate in the content validation stage of the development of a patellofemoral evaluation tool.

For those people who indicated that they do not see many patellofemoral patients, your contribution is still very valuable. By using feedback from a group of clinicians with expertise in the musculo-skeletal area but less exposure to a patellofemoral pain population, the content validation design will be stronger.

Enclosed is the first draft of the evaluation tool. Please fill in your name at the top of each section of the form and then read the instructions carefully before completing the questionnaire. Could you please return this questionnaire by **FEBRUARY 8** in the enclosed envelope.

Thank you very much for your help.

Yours sincerely

Liz Harrison, M.Sc., B.P.T.

LH/imr
encl.

INSTRUCTIONS AND BACKGROUND INFORMATION FOR COMPLETING THE ENCLOSED QUESTIONNAIRE

PLEASE READ CAREFULLY

Purpose: To develop an evaluation tool that can be used in clinical practices to monitor change in patients presenting with patellofemoral pain syndrome (PFPS).

Selection of Questions/Tests: The items listed have been selected based on existing evaluation systems for knee problems and considering preliminary analysis of data from a pilot study. Please note that questions appear in random order in order that reviewers consider each question on its own merit. In the final questionnaires, there will be a logical ordering of questions.

DEFINITIONS:

PFPS (Patellofemoral pain syndrome): There are many terms used to describe this syndrome including: patellagia, anterior knee pain, retropatellar arthralgia and patellar malalignment. For the purposes of this study, the operational definition of PFPS will include all of the following:

- Retropatellar or peripatellar pain;
- Presenting symptoms are related to the patella - patellofemoral articulation or the peripatellar structures; and
- Presenting symptoms may either be associated with articular surface changes, or as a result of patellar tracking problems.

Excluded from this syndrome are: peripatellar bursitis, tendinitis, plical syndromes, osteochondral fractures, osteochondritis dissecans, patellar subluxation or dislocation and vascular insufficiencies.

INSTRUCTIONS

There are two formats for responding to the following questions.

1. Put an "X" in the appropriate box; or

2. Mark your rating on a line, as indicated in the following example.

Example: The line is labelled at either end with descriptions related to the question. Place a mark at a point on the line that best describes your rating relative to the definers.

Not Important — "Mark" — Very Important

If you feel that a question needs modification, please suggest how you feel it should be modified and provide an explanation. Space has been provided after each question for comments or modifications.

PLEASE PUT YOUR NAME ON THE TOP OF EACH SECTION AND RETURN IN THE ENCLOSED ENVELOPE AS SOON AS POSSIBLE. THANK YOU!

PATELLOFEMORAL EVALUATION

REVIEWER'S NAME: _____

SECTION ONE:

I. Indicate on the lines below how you would rate the importance of each of these areas relative to the overall evaluation of PFPS patients.

1. Pain Assessment

Not Important ——— Very important

2. Disability Assessment (Functional Limitations)

Not Important ——— Very important

3. Activity Level (Recreation, Sport, Lifestyle)

Not Important ——— Very important

4. Objective Clinical Tests (ie. knee R.O.M.)

Not Important ——— Very important

5. Subjective Clinical Tests (ie. amount of swelling)

Not Important ——— Very important

11. Are there any other areas that you feel must be included in the evaluation? | No | Yes |

If yes, define the areas and explain:

REVIEWER'S NAME: _____

SECTION TWO:

The following questions are designed to be completed independently by the patient. A visual analogue scale would be used to rate responses. (This is the type of scale that you used in Section One. In the final questionnaire the line will be 10 cm. in length. It has been reduced in these samples in order to save space.)

Wording for these questions has been purposely kept short and simple, as it has been shown that questionnaires must be aimed at a reading level no higher than that of a 12 year old.

Indicate whether you feel the question is Needed, Needed with Modifications or Not Needed when assessing for PFPS. Put an X in the box that corresponds with your judgement. The "Not Needed" boxes have been shaded for ease of data entry and are not meant to influence your response. Please provide suggested modifications with any explanations in the space provided.

QUESTIONS

1. How much pain do you have in your knee:

At its worst?

Needed	Modify	Not Needed

Modification:

No Pain — Pain as severe as it could be

At its least?

Needed	Modify	Not Needed

Modification:

No Pain — Pain as severe as it could be

As it usually feels?

Needed	Modify	Not Needed

Modification:

No Pain — Pain as severe as it could be

2. How often do you have pain in your knee?

Needed ☐ Modify ☐ Not Needed ☐

Modification:

Never — Constant Pain, even at rest

3. How long does the pain in your knee last?

Needed ☐ Modify ☐ Not Needed ☐

Modification:

Immediately goes away — Constant pain

4. How does the pain in your knee affect your lifestyle?

Needed ☐ Modify ☐ Not Needed ☐

Modification:

No change — Severely affected

5. Has the pain in your knee changed since your first visit to the clinic?

Needed ☐ Modify ☐ Not Needed ☐

Modification:

No Pain — Pain Worse

6. Do you have any problems or any discomfort in your knee with the following activities?

a. walking a short distance (about a city block)

Needed ☐ Modify ☐ Not Needed ☐

Modification:

No Problem — Unable to Do

b. running a short distance, say 100 meters (about a city block)

No Problem — Unable to Do

Needed | Modify | Not Needed

Modification:

c. climbing up 4 flights of stairs (about 32 steps)

No Problem — Unable to Do

Needed | Modify | Not Needed

Modification:

d. sitting for prolonged periods with your knees bent in one position

No Problem — Unable to Do

Needed | Modify | Not Needed

Modification:

e. kneeling

No Problem — Unable to Do

Needed | Modify | Not Needed

Modification:

f. squatting

No Problem — Unable to Do

Needed | Modify | Not Needed

Modification:

g. climbing up 2 flights of stairs (about 16 steps)

No Problem — Unable to Do

Needed | Modify | Not Needed

Modification:

h. walking as far as a mile

No Problem — Unable to Do

Needed | Modify | Not Needed

Modification:

i. jumping

Needed Modify Not Needed

No Problem — Unable to Do

Modification:

j. sitting for a short period with your knees bent in one position

Needed Modify Not Needed

No Problem — Unable to Do

Modification:

k. walking down 1 flight of stairs

Needed Modify Not Needed

No Problem — Unable to Do

Modification:

l. running as far as a mile

Needed Modify Not Needed

No Problem — Unable to Do

Modification:

7. How does your knee affect your normal activities?

Needed Modify Not Needed

No Limitations — Severely Limited

Modification:

8. What was your sports or physical activity level before your knee condition?

Needed Modify Not Needed

Participated in competitive sports or intense physical activity 5-7times/wk. — No sports or physical activities

Modification:

9. At the present time, what is your level of sports or recreational activity?

Needed ☐ Modify ☐ Not Needed ☐

Modification:

Participate in competitive sports or intense physical activity 5–7times/wk. — No sports or physical activities

10. Has your sports or recreational activity level changed due to your knee condition?

Needed ☐ Modify ☐ Not Needed ☐

Modification:

No Change — Given up all sports

11. Does your knee condition affect your performance in sport activities?

Needed ☐ Modify ☐ Not Needed ☐

Modification:

No Change — Severely limited

12. Has your performance at work changed due to your knee condition?

Needed ☐ Modify ☐ Not Needed ☐

Modification:

No Change — Unable to Work

COMMENTS:

REVIEWER'S NAME: _____

SECTION THREE:

The following tests/observations are frequently performed and recorded by clinicians. On the lines below, indicate your rating of the importance of each test, considering the ability of the test to **DETECT CHANGES** **relative to the patient's condition.**

Assume standardized test procedures and operational definitions for measurements are used.

1. Amount of swelling in the knee.

Not Important _____ Very Important

2. Amount of crepitus at the patellofemoral joint.

Not Important _____ Very Important

3. Muscle tone of the quadriceps muscle group during isometric contraction.

Not Important _____ Very Important

4. The degree of pain during Clarke's compression test.

Not Important _____ Very Important

5. The amount of pain during passive patellar movements.

Not Important — Very Important

6. The amount of passive mobility of the patella.

Not Important — Very Important

7. The flexibility of the hamstring muscle group.

Not Important — Very Important

8. The flexibility of the rectus femoris muscle.

Not Important — Very Important

9. Flexibility of the hip flexors.

Not Important — Very Important

10. Flexibility of tensor fascia latae.

Not Important — Very Important

11. Flexibility of adductor muscle group.

Not Important — Very Important

12. Flexibility of gastrocnemius and soleus muscle groups.

Not Important — Very Important

13. Static patellar alignment (orientation).

Not Important — Very Important

14. Lower extremity alignment in standing.

Not Important — Very Important

15. Limb girth measured at equal increments from knee joint line to mid-thigh.

Not Important — Very Important

16. Gait evaluation.

Not Important — Very Important

17. Time to onset of pain during a step test.

Not Important ——— Very Important

18. Q-angle.

Not Important ——— Very Important

19. Knee range of motion.

Not Important ——— Very Important

20. Radiographs.

Not Important ——— Very Important

21. Isokinetic muscle evaluation of quadriceps and hamstrings.

Not Important ——— Very Important

COMMENTS:

APPENDIX E

CLINICAL EVALUATION FORMS

INITIAL AND FINAL EVALUATIONS

04/93

PATELLOFEMORAL INITIAL EVALUATION

SUBJECT #

HISTORY

ID NUMBER | LAST NAME | FIRST NAME & INITIAL | SEX M [ ] F [ ] | AGE YRS | HT. CM | WT. KG

DATE FIRST VISIT DAY MO YR

KNEE INVOLVED
R [ ] BILAT WITH R > L [ ]
L [ ] BILAT WITH L > R [ ]

ONSET OF INITIAL PROBLEM
< 6 WKS [ ] 6 – 12 WKS [ ] 6 – 12 MOS [ ] 3 – 6 YRS [ ]
3 – 6 WKS [ ] 3 – 6 MOS [ ] 1 – 3 YRS [ ] > 6 YRS [ ]

ETIOLOGY OF CONDITION
NO KNOWN CAUSE [ ]
OVERUSE [ ]
ACUTE [ ]
COMMENTS:

PREVIOUS KNEE SURGERY

AFFECTED SIDE	NON-AFFECTED SIDE
NO [ ]	NO [ ]
YES [ ]	YES [ ]

SPECIFY:

PREVIOUS KNEE INJURY
REQUIRED MEDICAL ATTENTION OR LIMITED ADL > 2 DAYS

AFFECTED SIDE	NON-AFFECTED SIDE
NO [ ]	NO [ ]
YES [ ]	YES [ ]

SPECIFY:

MEDICAL HISTORY

	YES	NO		
LUMBO-SACRAL	[ ]	[ ]	[ ]	[ ]
HIP	[ ]	[ ]	[ ]	[ ]
ANKLE	[ ]	[ ]	[ ]	[ ]

MEDICATIONS:
COMMENTS:

PREVIOUS TREATMENT OF PATELLOFEMORAL PROBLEM

NONE [ ]
MEDICAL [ ]
PHYSICAL THERAPY [ ]
OTHER [ ]
SPECIFY:

MANAGEMENT
MEDICATIONS [ ] SUPERVISED PROGRAM [ ] ORTHOTICS [ ]
HOME PROGRAM [ ] EDUCATION [ ] TAPING [ ]
OTHER [ ] SPECIFY:

PREVIOUS OUTCOME
IMPROVED [ ]
NO CHANGE [ ]
WORSE [ ]
N/A [ ]

ACTIVITY RATING

BASED ON HISTORY OF PHYSICAL ACTIVITY RATE THE OVERALL ACTIVITY LEVEL

HIGH — MODERATE — LOW — VERY LOW

PHYSICAL EVALUATION

LOWER QUADRANT SCAN
NEGATIVE [ ]
POSITIVE [ ]
Note: Lower Limb alignment not considered

IF POSITIVE SPECIFY: NEUROLOGICAL [ ] SPINE [ ] PELVIS [ ]

COMMENTS:

LOWER EXTREMITY ALIGNMENT

INDICATE RATING ON VISUAL ANALOGUE SCALE

RIGHT: NORMAL — MINIMAL CHANGE — MODERATE CHANGE — SEVERE ABNORMALITY

LEFT: NORMAL — MINIMAL CHANGE — MODERATE CHANGE — SEVERE ABNORMALITY

HIP	R	L
NORMAL		
ANTEVERT		
RETROVERT		

KNEE	R	L
NORMAL		
GENU VALGUM		
GENU VARUM		
HYPEREXT		
EXT TIB TORSION		
INT TIB TORSION		

PFJ	R	L
NORMAL		
MEDIAL		
LATERAL		

FOOT / ANKLE	R	L
NORMAL		
CALC VALGUS		
CALC VARUS		
FF VALGUS		
FF VARUS		
FLAT ARCH		
PES CAVUS		

LEG LENGTH	
NORMAL	
< 2 CM	
> 2 CM	

PASSIVE PATELLAR MOBILITY

INDICATE RATING ON VISUAL ANALOGUE SCALE

RIGHT: HYPERMOBILE — NORMAL — HYPOMOBILE

LEFT: HYPERMOBILE — NORMAL — HYPOMOBILE

STATIC PATELLAR ORIENTATION

INDICATE RATING ON VISUAL ANALOGUE SCALE

RIGHT

NORMAL — MINIMAL CHANGE — MODERATE CHANGE — SEVERE ABNORMALITY

LEFT

NORMAL — MINIMAL CHANGE — MODERATE CHANGE — SEVERE ABNORMALITY

ROTATION	R	L	TILT	R	L	SUPERIOR/INFERIOR	R	L
Inferior Border			Medial Border					
NORMAL			NORMAL			NORMAL		
INTERNAL			ANTERIOR			SUPERIOR (ALTA)		
EXTERNAL			POSTERIOR			INFERIOR (BAJA)		

FLEXIBILITY RECTUS FEMORIS

INDICATE RATING ON VISUAL ANALOGUE SCALE

RIGHT

NORMAL — MINIMAL DECREASE — MODERATE DECREASE — MARKED DECREASE

LEFT

NORMAL — MINIMAL DECREASE — MODERATE DECREASE — MARKED DECREASE

FLEXIBILITY TENSOR FASCIA LATAE

INDICATE RATING ON VISUAL ANALOGUE SCALE

RIGHT

NORMAL — MINIMAL DECREASE — MODERATE DECREASE — MARKED DECREASE

LEFT

NORMAL — MINIMAL DECREASE — MODERATE DECREASE — MARKED DECREASE

PAIN RATING

BASED ON PAIN DURING ASSESSMENT INDICATE AN OVERALL PAIN RATING

NO PAIN — MODERATE — SEVERE PAIN

FUNCTION RATING

BASED ON FUNCTIONAL TESTING INDICATE AN OVERALL FUNCTION RATING

NORMAL — MODERATELY LIMITED — SEVERELY LIMITED

OTHER FINDINGS:

ASSESSMENT COMPLETED ON ONE VISIT YES [ ] NO [ ]

PROGNOSIS

BASED ON ASSESSMENT FINDINGS INDICATE A RATING FOR THIS PATIENT'S PROGNOSIS OVER A ONE MONTH PERIOD

CONDITION WORSE — NO CHANGE — COMPLETE RECOVERY

TREATMENT PLAN:

EVALUATORS NAME _____ *PT* [ ] *MD* [ ]

1/90

06/93 Date: **PATELLOFEMORAL FINAL EVALUATION** 194 SUBJECT #

ID NUMBER | LAST NAME | FIRST NAME & INITIAL | KNEE INVOLVED

R [ ] BILAT WITH R > L [ ]
L [ ] BILAT WITH L > R [ ]

TREATMENT PROGRAM GIVEN FOR PATELLOFEMORAL PAIN PROBLEM

NONE [ ]
MEDICAL [ ]
PHYSICAL THERAPY [ ]
OTHER [ ]
SPECIFY:

MANAGEMENT
MEDICATIONS [ ] SUPERVISED PROGRAM [ ] ORTHOTICS [ ]
HOME PROGRAM [ ] EDUCATION [ ] TAPING [ ]
OTHER [ ] SPECIFY:

COMMENTS:

TREATMENT COMPLIANCE

POOR [ ]
FAIR [ ]
GOOD [ ]
COMMENTS:

TIME FRAME SINCE INITIAL VISIT: [ ] < 3 WKS. [ ] 3 TO 6 WKS. [ ] > 6 WKS.

ACTIVITY RATING

BASED ON HISTORY OF PHYSICAL ACTIVITY RATE THE OVERALL ACTIVITY LEVEL

HIGH — MODERATE — LOW — VERY LOW

PHYSICAL EVALUATION

LOWER QUADRANT SCAN

NEGATIVE [ ]
POSITIVE [ ]
Note: Lower Limb alignment not considered

IF POSITIVE SPECIFY: NEUROLOGICAL [ ] SPINE [ ] PELVIS [ ]
COMMENTS:

LOWER EXTREMITY ALIGNMENT

INDICATE RATING ON VISUAL ANALOGUE SCALE

RIGHT: NORMAL — MINIMAL CHANGE — MODERATE CHANGE — SEVERE ABNORMALITY

LEFT: NORMAL — MINIMAL CHANGE — MODERATE CHANGE — SEVERE ABNORMALITY

HIP	R	L
NORMAL		
ANTEVERT		
RETROVERT		

KNEE	R	L
NORMAL		
GENU VALGUM		
GENU VARUM		
HYPEREXT		
EXT TIB TORSION		
INT TIB TORSION		

PFJ	R	L
NORMAL		
MEDIAL		
LATERAL		

FOOT / ANKLE	R	L
NORMAL		
CALC VALGUS		
CALC VARUS		
FF VALGUS		
FF VARUS		
FLAT ARCH		
PES CAVUS		

LEG LENGTH	
NORMAL	
< 2 CM	
> 2 CM	

STATIC PATELLAR ORIENTATION

INDICATE RATING ON VISUAL ANALOGUE SCALE

RIGHT: NORMAL — MINIMAL CHANGE — MODERATE CHANGE — SEVERE ABNORMALITY

LEFT: NORMAL — MINIMAL CHANGE — MODERATE CHANGE — SEVERE ABNORMALITY

ROTATION (Inferior Border)	R	L
NORMAL		
INTERNAL		
EXTERNAL		

TILT (Medial Border)	R	L
NORMAL		
ANTERIOR		
POSTERIOR		

SUPERIOR/INFERIOR	R	L
NORMAL		
SUPERIOR (ALTA)		
INFERIOR (BAJA)		

PASSIVE PATELLAR MOBILITY

RIGHT

INDICATE RATING ON VISUAL ANALOGUE SCALE

HYPERMOBILE — NORMAL — HYPOMOBILE

LEFT

HYPERMOBILE — NORMAL — HYPOMOBILE

FLEXIBILITY RECTUS FEMORIS

RIGHT

INDICATE RATING ON VISUAL ANALOGUE SCALE

NORMAL — MINIMAL DECREASE — MODERATE DECREASE — MARKED DECREASE

LEFT

NORMAL — MINIMAL DECREASE — MODERATE DECREASE — MARKED DECREASE

FLEXIBILITY TENSOR FASCIA LATAE

RIGHT

INDICATE RATING ON VISUAL ANALOGUE SCALE

NORMAL — MINIMAL DECREASE — MODERATE DECREASE — MARKED DECREASE

LEFT

NORMAL — MINIMAL DECREASE — MODERATE DECREASE — MARKED DECREASE

PAIN RATING

BASED ON PAIN DURING CLINICAL TESTING INDICATE AN OVERALL PAIN RATING

NO PAIN — MODERATE — SEVERE PAIN

FUNCTION RATING

BASED ON FUNCTIONAL TESTING INDICATE AN OVERALL FUNCTION RATING

NORMAL — MODERATELY LIMITED — SEVERELY LIMITED

OTHER FINDINGS:

REASSESSMENT COMPLETED ON ONE VISIT YES [ ] NO [ ]

CLINICAL CHANGE

RATE THIS PATIENT'S CHANGE SINCE INITIAL ASSESSMENT

CONDITION WORSE — NO CHANGE — COMPLETE RECOVERY

PLAN: D/C [ ] CONTINUE TREATMENT [ ] CHANGE TREATMENT [ ] FURTHER REFERRAL [ ]

COMMENTS:

EVALUATORS NAME _____ *PT [ ] MD [ ]*

06/93

APPENDIX F

PATIENT QUESTIONNAIRES

INITIAL AND FINAL QUESTIONNAIRES

INITIAL QUESTIONAIRRE | *NAME:* | *DATE:*

Give the best answer to each of the following questions based on your knee condition today. Mark a point on each of the lines below that best describes your knee condition, relative to the words along or at the ends of the line.

The following is an example of how to mark the line:

1. Rate your pain at its worst:

2. Rate your pain at its least:

3. Rate your pain as it usually feels:

4. How often do you have pain in your knee?

5. How long until the pain in your knee goes away after activity?

6. How does the pain in your knee affect your daily activities?

7. Do you have any problems or pain in your knee with the following activities?

a. walking a short distance (about a city block)

b. running a short distance (about a city block)

c. walking as far as a mile

d. running as far as a mile

e. climbing up 1 flight of stairs (about 8 steps)

f. climbing up 4 flights of stairs (about 32 steps)

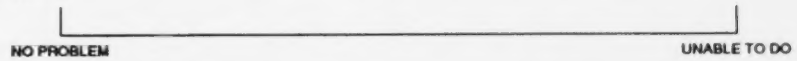
g. climbing down 1 flight of stairs (about 8 steps)

h. climbing down 4 flights of stairs (about 32 steps)

i. sitting for 15 minutes with your knees bent in one position

j. sitting for 1 hour with your knees bent in one position

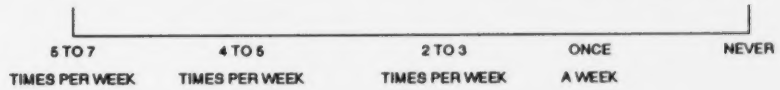
k. kneeling



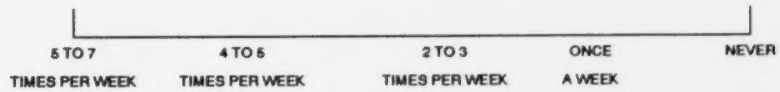
l. squatting

m. jumping

8. How often did you participate in intense physical activity before your knee condition?



9. Presently, how often do you participate in intense physical activity?



10. How does your knee condition affect your physical activity level?

11. How has your performance at work changed due to your knee condition?

Circle the best answer for questions 12 to 15.

12. Thinking about the things you do at work, how would you rate yourself as to the amount of physical activity you get compared with others of your age and sex?
    a. Much more active
    b. Somewhat more active
    c. About the same
    d. Somewhat less active
    e. Much less active
    f. Not applicable

13. Now, thinking about the things you do outside of work, how would you rate yourself as to the amount of physical activity you get compared with others your age and sex?
    a. Much more active
    b. Somewhat more active
    c. About the same
    d. Somewhat less active
    e. Much less active

14. Do you regularly engage in strenuous exercise or hard physical labor?
    a. Yes (Answer Question #16)
    b. No (Stop)

15. Do you exercise or labor at least three times a week?
    a. Yes
    b. No

16. Considering a typical day during the work or school week check the box which best describes how much time you spend doing each of these activities over 24 hours.

a. SITTING

☐ 7 - 8 hrs./day ☐ 5 - 6 hrs./day ☐ 3 - 4 hrs./day ☐ 1 - 2 hrs./day ☐ less than 1 hr./day

b. LYING

☐ 9 - 10 hrs./day ☐ 7 - 8 hrs./day ☐ 5 - 6 hrs./day ☐ 3 - 4 hrs./day ☐ 1 - 2 hrs./day

c. STANDING

☐ 7 - 8 hrs./day ☐ 5 - 6 hrs./day ☐ 3 - 4 hrs./day ☐ 1 - 2 hrs./day ☐ less than 1 hr./day

d. WALKING

☐ 7 - 8 hrs./day ☐ 5 - 6 hrs./day ☐ 3 - 4 hrs./day ☐ 1 - 2 hrs./day ☐ less than 1 hr./day

e. CLIMBING

☐ 4 flights 2 times/day ☐ 3 flights 2 times/day ☐ 2 flights 2 times/day ☐ 1 flight 2 times/day ☐ 0 flights /day

f. SQUATTING OR KNEELING

☐ More than 16 times/day ☐ 11 - 15 times/day ☐ 6 - 10 times/day ☐ 1 - 5 times/day ☐ 0 times/day

g. CARRYING OR LIFTING

☐ More than 16 times/day ☐ 11 - 15 times/day ☐ 6 - 10 times/day ☐ 1 - 5 times/day ☐ 0 times/day

h. SPORTS OR RECREATIONAL ACTIVITIES

☐ 2 hrs. or more/day ☐ 90 minutes /day ☐ 1 hr. /day ☐ 30 minutes /day ☐ 0 minutes /day

FINAL QUESTIONAIRRE | *NAME:* | *DATE:*

Give the best answer to each of the following questions based on your knee condition today. Mark a point on each of the lines below that best describes your knee condition, relative to the words along or at the ends of the line.

The following is an example of how to mark the line:

NO PAIN — PAIN AS SEVERE AS IT COULD BE

1. Rate your pain at its worst:

NO PAIN — PAIN AS SEVERE AS IT COULD BE

2. Rate your pain at its least:

NO PAIN — PAIN AS SEVERE AS IT COULD BE

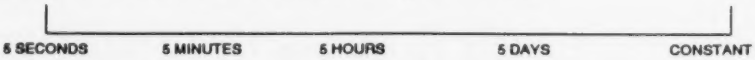
3. Rate your pain as it usually feels:

NO PAIN — PAIN AS SEVERE AS IT COULD BE

4. How often do you have pain in your knee?

NEVER — ONCE A WEEK — ONCE A DAY — ONCE AN HOUR — CONSTANTLY

5. How long until the pain in your knee goes away after activity?



5 SECONDS — 5 MINUTES — 5 HOURS — 5 DAYS — CONSTANT PAIN

6. How does the pain in your knee affect your daily activities?

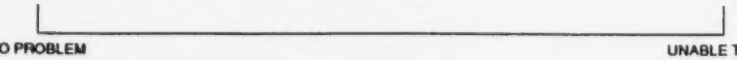
NO CHANGE — SEVERELY LIMITED

7. Do you have any problems or pain in your knee with the following activities?

a. walking a short distance (about a city block)

NO PROBLEM — TO DO

b. running a short distance (about a city block)



NO PROBLEM — UNABLE TO DO

c. walking as far as a mile

NO PROBLEM — UNABLE TO DO

d. running as far as a mile

e. climbing up 1 flight of stairs (about 8 steps)

f. climbing up 4 flights of stairs (about 32 steps)

g. climbing down 1 flight of stairs (about 8 steps)

h. climbing down 4 flights of stairs (about 32 steps)

i. sitting for 15 minutes with your knees bent in one position

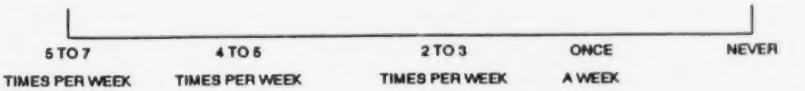
j. sitting for 1 hour with your knees bent in one position

k. kneeling

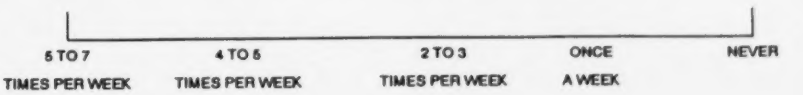
l. squatting

m. jumping

8. How often did you participate in intense physical activity before your knee condition?



9. Presently, how often do you participate in intense physical activity?



10. How does your knee condition affect your physical activity level?

11. How has your performance at work changed due to your knee condition?

Circle the best answer for questions 12 to 15.

12. Thinking about the things you do at work, how would you rate yourself as to the amount of physical activity you get compared with others of your age and sex?
    a. Much more active
    b. Somewhat more active
    c. About the same
    d. Somewhat less active
    e. Much less active
    f. Not applicable

13. Now, thinking about the things you do outside of work, how would you rate yourself as to the amount of physical activity you get compared with others your age and sex?
    a. Much more active
    b. Somewhat more active
    c. About the same
    d. Somewhat less active
    e. Much less active

14. Do you regularly engage in strenuous exercise or hard physical labor?
    a. Yes (Answer Question #16)
    b. No (Stop)

15. Do you exercise or labor at least three times a week?
    a. Yes
    b. No

16. Considering a typical day during the work or school week check the box which best describes how much time you spend doing each of these activities over 24 hours.

a. SITTING

☐ 7 - 8 hrs./day ☐ 5 - 6 hrs./day ☐ 3 - 4 hrs./day ☐ 1 - 2 hrs./day ☐ less than 1 hr./day

b. LYING

☐ 9 - 10 hrs./day ☐ 7 - 8 hrs./day ☐ 5 - 6 hrs./day ☐ 3 - 4 hrs./day ☐ 1 - 2 hrs./day

c. STANDING

☐ 7 - 8 hrs./day ☐ 5 - 6 hrs./day ☐ 3 - 4 hrs./day ☐ 1 - 2 hrs./day ☐ less than 1 hr./day

d. WALKING

☐ 7 - 8 hrs./day ☐ 5 - 6 hrs./day ☐ 3 - 4 hrs./day ☐ 1 - 2 hrs./day ☐ less than 1 hr./day

e. CLIMBING

☐ 4 flights 2 times/day ☐ 3 flights 2 times/day ☐ 2 flights 2 times/day ☐ 1 flight 2 times/day ☐ 0 flights /day

f. SQUATTING OR KNEELING

☐ More than 16 times/day ☐ 11 - 15 times/day ☐ 6 - 10 times/day ☐ 1 - 5 times/day ☐ 0 times/day

g. CARRYING OR LIFTING

☐ More than 16 times/day ☐ 11 - 15 times/day ☐ 6 - 10 times/day ☐ 1 - 5 times/day ☐ 0 times/day

h. SPORTS OR RECREATIONAL ACTIVITIES

☐ 2 hrs. or more/day ☐ 90 minutes /day ☐ 1 hr. /day ☐ 30 minutes /day ☐ 0 minutes /day

APPENDIX G

OPERATIONAL MANUAL

TO: Physical Therapists

FROM: Liz Harrison

JUNE 1993

Thankyou for agreeing to participate in the Patellofemoral Pain Syndrome Evaluation study. Enclosed are documents reviewing the inclusion/exclusion criteria for patients, the methodology for completing the evaluation forms and patient questionnaires, operational definitions for each of the areas of the evaluation form and the forms.

As you know standardization of evaluation is essential for measures to be objective. Therefore, I have provided a definition for each area of the evaluation and hope that you will be able to follow these definitions when completing the form. I have piloted the forms on a group of clinicians with limited experience in evaluating patellofemoral patients and they found that they could carry out the evaluation and complete the form in under 30 minutes. I am hoping that you will be able to easily integrate this assessment into your usual clinical setting.

Please fill in all areas of the form at the initial and final visits. After completing each evaluation form (initial and final) please make a copy of the form and file it in the folder provided. I have also piloted the patient questionnaires and each patient took less than 10 minutes to complete the questionnaire. The questionnaires are extremely important in this study, so please ensure that each patient completes a form at the initial and final visits. Again please make copies of the forms and file them along with the evaluation forms. Please ensure that patient identification appears on each form. The data will be coded and names removed from the forms on receipt, to ensure patient confidentiality.

If you have any questions please do not hesitate to call me or FAX me. (Phone: Office (306) 966-6579 Home (306) 343-1091 FAX (306) 966-6575)

Thankyou again for your help! We need to develop accurate methods to evaluate outcomes in our patients and your assistance in this study will be a step towards improving clinical measurement.

Sincerely,

Liz Harrison

STANDARDIZED METHODOLOGY FOR

PATELLOFEMORAL PAIN SYNDROME EVALUATION STUDY

STAGES OF STUDY

INCLUSION/EXCLUSION CRITERIA

PATIENT QUESTIONNAIRE

CLINICAL EVALUATION

OPERATIONAL DEFINITIONS

FINAL EVALUATION FORM

STAGES FOR EVALUATION STUDY OF PATELLOFEMORAL PAIN SYNDROME PATIENTS

1. Ensure patient meets inclusion/exclusion criteria.

2. Have patient review and sign consent form.

3. Ensure that patient name or identification number is filled in on evaluation form and questionnaire.

4. Perform evaluation and have patient fill in questionnaire (in order to standardize the process have patient fill in questionnaire PRIOR to the clinical evaluation).

5. Perform clinical evaluation and complete form.

6. File copies of the evaluation form and patient questionnaire in research folder.

PFPS INCLUSION AND EXCLUSION CRITERIA

INCLUSION:

Age group: 14 to 45 year olds, males and females
Activity Level: At minimum, subject must be independent in basic ADL and ambulating independently (ie. no external support)
History - pt. may or may not identify an initial episode or mechanism of injury
Must present with the following:

Patellofemoral pain aggravated with activity
Pain present with palpation in the patellar region (ie. medial or lateral patellar facets or femoral condyles) Note: if only inferior fat pad involved subject excluded (however, if the inferior fat pad is involved in addition to the Patellofemoral joint then the patient would be included)
Patellofemoral pain present with resisted knee extension (either dynamic or isometric)

EXCLUSION:

-Positive neurological signs/symptoms referred from lumbo-sacral region
-Recent injury (last 6 months) to other lower extremity joints or spine/pelvis that affected function for more than 2 days
-Knee injury including: any ligamentous or meniscal injury that has been surgically treated; ligamentous or meniscal injury that has been medically diagnosed and/or treated in the past 12 months; any ligamentous or meniscal injury that has been medically diagnosed and treated resulting in more than 3 months of restricted activity
-Extensor apparatus involvement (quadriceps or patellar tendinitis)
-Bursitis
-Isolated inferior fat pad involvement
-Reflex sympathetic dystrophy
-Radiological changes associated with osteoarthritis, osteochondral or chondral fractures of the knee
-Dislocation or subluxation of the patella
-Recent steroid (one year) injections to the knee
-Gross knee effusion

PATIENT QUESTIONNAIRE

1. Have each patient complete the questionnaire PRIOR to you performing the clinical evaluation at initial and final sessions. By completing the questionnaire prior to evaluation, bias will be reduced. (ie. interaction with therapist, evaluation process)

2. Patients must be able to read and understand the questions. The questions are a combination of visual analogue scales (which require patient to put a mark at each spot on the line) and circle the best answer. Although the instructions for completing the questionnaire are on the top of the form, please make sure on the initial visit that you review the method of completing the form with the patient.

CLINICAL EVALUATION

1. The same therapist should evaluate and re-evaluate patients.

2. Ensure that all forms have patient identification and date on them.

3. A final evaluation form should be completed at final visit (when treatment discontinued) or at one month following initial visit. If patient is treated for more than one month, fill in an additional Final Evaluation form when the patient is discontinued from treatment. The purpose of this study is to evaluate the effectiveness of the form for detecting change, thus the one month period was considered a useful time frame for re-evaluation. In some cases patients may be treated for a shorter period of time, thus a final evaluation should be performed at completion of treatment period.

4. Every attempt should be made to complete all aspects of the evaluation (initial and final) and questionnaires in one session. A spot on the evaluation form has been included to indicate that evaluation has been completed in one visit. If more than one session is needed to complete the evaluation, error will be introduced into the study.

5. All patients should have the following filed: initial evaluation, initial questionnaire, final evaluation, final questionnaire and consent form.

OPERATIONAL DEFINITIONS FOR EVALUATION FORM

Note: Except for the History section of the evaluation form, the majority of the recording is to be done on Visual Analogue Scales. Space is provided for additional notes, use as appropriate. For the purposes of this study, additional detail will be used primarily for clarification.

Operational definitions are provided for each component of the evaluation form in order to improve standardization, as different therapists will be using the form. Although there may be different interpretations for each scale, please attempt to use the operational definitions as basis for decision making. It is assumed that the operational definitions will help control for bias and variations in clinical interpretation. However, it is also recognized that subjective judgement makes up clinical testing and that this will be reflected in the evaluation. Clinicians should use the operational definitions as guidelines, and adapt these to deal with each individual.

HISTORY - put an X in the appropriate box

COMPONENTS:	
ID NUMBER	
NAME	
SEX	
AGE	In years
HEIGHT	In centimetres
WEIGHT	In kilograms
DATE OF FIRST VISIT	Patient's initial visit to you for this problem
KNEE INVOLVED	If bilateral indicate which knee is more involved
ONSET OF INITIAL PROBLEM	First signs/symptoms
ETIOLOGY	Comments optional but may be helpful if specific mechanism noted
PREVIOUS KNEE SURGERY	Indicate procedure reported
PREVIOUS KNEE INJURY	Consider a knee injury as significant if medical attention was required or if subject reported limited activity (ie. ambulation) for more than 2 days. Indicate type of injury.
MEDICAL HISTORY	Indicate if the subject reports any medical problem in other regions that required medical attention or have affected function (ie. occasional back pain would not be considered a positive history, unless accompanied by altered function for a period of time or medical attention)
PREVIOUS TREATMENT	Indicate treatment for Patellofemoral problem only
PREVIOUS OUTCOME	Base this on patient's judgement

ACTIVITY RATING

The following guidelines should be used in order for the clinician to rate the patient's activity level. Consider physical activity as recreational or competitive sports, leisure activities and occupation (labour). The rating should be based on patient's current activity (ie. past few weeks).

HIGH: Subject participates in intense physical activity 5 to 7 times a week for a minimum of 1 hour. The clinician must evaluate the intensity of the activity as well as the duration in making this judgement.

MODERATE: Subject participates in intense physical activity or labour 3 to 5 times a week for a minimum of 45 minutes. The clinician may judge less intense activities (ie. casual walk) but with longer duration (ie. greater than 1 hour) in this category. However the frequency should be maintained at 3 to 5 times per week.

LOW: Subject participates in intense physical activity one or two times a week but not regularly. Infrequent activity will be more commonly found in this category. Length of activity will vary, but usually less than 30 minutes.

VERY LOW: Subject does not participate in physical activity at work/school or outside work/school.

CLINICAL TESTS

LOWER QUADRANT SCAN

COMPONENTS:

1. POSTURE: (anterior, posterior, and saggital views)
   - Spine
   - Pelvis

NOTE: Lower extremity alignment not to be considered in this scan as it will be covered in detail later in the evaluation. Therefore do not indicate a positive scan if only lower extremity alignment is abnormal.

2. NEUROLOGICAL:
   - Lower Extremity Motor and Sensory Function

3. SIGNS/SYMPTOMS PRESENT:

   SPINE, PELVIS, HIP AND ANKLE

NOTE: Knee signs and symptoms not to be considered in this scan. If only knee signs and/or symptoms are present do not indicate a positive scan.

NEGATIVE SCAN:

Normal posture of spine and pelvis
No neurological findings
No signs or symptoms in spine, pelvis, hip and ankle

POSITIVE SCAN:
At least one postural component abnormal
OR
Neurological findings (motor and/or sensory)
OR
Signs or symptoms in spine, pelvis, hip or ankle

LOWER EXTREMITY ALIGNMENT

Evaluation of the lower limb in relaxed standing, with the feet shoulder width apart.

Provide an overall rating for each leg and indicate specific findings by filling in the appropriate boxes.

Components:	Clinical Measurement
1. HIP Femoral head/neck/shaft position	NORMAL: Greater trochanter below mid-point of Nelaton's line (ASIS to ischial tuberosity) ABNORMAL:Anteversion: greater trochanter anterior to mid-point Retroversion: greater trochanter posterior to mid-point
2. KNEE Tibio-femoral angle frontal	NORMAL: Tibio-femoral angle 6 degrees varus ABNORMAL: Genu valgum - 3 cm. or more between medial malleoli Genu varum - 3 cm. or more between medial femoral condyles
3. KNEE Tibial torsion	NORMAL: Tibial spine straight, no bowing; medial malleoli slightly anterior to lateral malleoli ABNORMAL: Tibial bowing and/or medial malleoli posterior to lateral malleoli or more than 10 degrees externally rotated from neutral position
4. KNEE Sagittal Plane position	NORMAL: Knee 0 to 5 degrees of knee flexion ABNORMAL: Knee hyperextension or > 5 degrees of knee flexion
5. PATELLOFEMORAL Anterior	NORMAL: Patella aligned between the 2nd and 3rd toes. ABNORMAL: Patella aligned with the 1st toe or the 4th or 5th toes
6. FOOT/ANKLE Tibio-Calcaneal Angle	NORMAL: Bisected calcaneal vertical line aligns with bisected lower one-third of lower leg ABNORMAL: Rearfoot valgus - > 5 degrees calcaneal valgus Rearfoot varus - < 5 degrees of calcaneal varus

7. FOOT/ANKLE Forefoot Angle	NORMAL: Plantar surfaces of rearfoot and forefoot in contact with supporting surface and parallel in alignment, with all metatarsals bearing weight ABNORMAL: Forefoot varus - medial forefoot decreased contact with supporting surface and/or increased weight bearing lateral metatarsal heads Forefoot valgus - lateral forefoot decreased contact with supporting surface and/or increased weight bearing medial metatarsal heads
8. FOOT/ANKLE Medial Arch	NORMAL: Medial longitudinal arch contour observable and medial forefoot (navicular region) not in contact with the supporting surface ABNORMAL: Pes planus (flat) - medial arch contour not observable OR medial forefoot in contact with supporting surface Pes cavus(high)- medial arch exaggerated with inability to flatten arch when hip and leg externally rotated
LEG LENGTH	NORMAL: Iliac crests level and Legs equal in length ABNORMAL: Leg differ by > 2 cm (ASIS to lat. malleoli)

NORMAL: (Rate each leg individually)
Hip, knee, patellofemoral and foot/ankle alignments within normal standards.
ABNORMAL:
Minimal: - less than 3 components demonstrate abnormality AND abnormality corrects with minimal voluntary effort from patient

Moderate: - less than 3 components abnormal but abnormality corrects with much voluntary effort from the patient or external support

OR

- 3 to 5 components demonstrate abnormality AND abnormality corrects with minimal effort

OR

- leg length differs by > 2 cm

Severe: - 5 to 8 components demonstrate abnormality
OR
- unable to correct any abnormality

PASSIVE PATELLAR MOBILITY

With the patient relaxed in supine position. Provide a rating for each patella individually.

Passive movements in medial and lateral directions (frontal plane) and superior and inferior directions (frontal plane) are performed.

HYPERMOBILITY: Passive movements in all directions are beyond normal limits with decreased tissue tension palpated.

NORMAL: Passive movements in all direction are within normal limits with reasonable tissue tension palpated at the end range of each movement.

HYPOMOBILITY: Passive movements in all directions are restricted with increased tissue tension palpated.

STATIC PATELLAR ALIGNMENT

In supine with legs relaxed and in normal resting position. Provide an overall rating for each patella individually and indicate specific findings by checking appropriate boxes.

NORMAL:

Patella mid-line, sitting above the trochlear groove with inferior border in contact with suprapatellar fat pad. The medial border and lateral border or the patella are level (ie. parallel with the frontal plane) and the superior border and inferior border are vertically aligned.

ABNORMAL:
Minimal: - Medial border and lateral border of patella are not level
OR
- Superior border and inferior border are not aligned vertically
OR
- Patella sitting superior or inferior to normal position by less than 1 cm

Moderate: - Any two of the above findings
OR

- Patella baja or alta

Severe: - All components mentioned in minimal
OR
- Patella baja or alta in combination with any of the other findings

FLEXIBILITY OF RECTUS FEMORIS

Testing position: Thomas test position. Therapist supporting non-test leg in full hip flexion and ensuring lumbar spine flat and pelvis stable. Provide a rating for each leg individually.

NORMAL:
Hip resting in 0 degrees flexion, lumbar spine straight (no lordosis) and knee flexed to 90 degrees. No compensation in lumbar spine or pelvis.

ABNORMAL:
Minimal: Hip neutral, knee flexion 70 to 90 degrees
Moderate: Hip neutral, knee flexion 50 to 70 degrees
Severe: Hip neutral, knee flexion less than 50 degrees

If hip flexion present then subtract 10 degrees of knee flexion from each category (normal=knee flex. 80 degrees, minimal=knee flex. 50 to 70 degrees, moderate=knee flex. 40 to 60 degrees, severe=knee flex. <40 degrees)

FLEXIBILITY OF TENSOR FASCIAE LATAE

Testing position: Thomas test position as above. Provide a rating for each leg individually.

NORMAL:
Hip positioned in mid-position (neutral abduction/adduction and neutral rotation) and lumbar spine straight (no lordosis). Pelvis stable.

ABNORMAL:
Minimal: Hip held in < 10 degrees abduction or < 10 degrees int. rotation but able to passively adduct hip to neutral position without pelvic movement.

Moderate: Hip held in > 10 degrees abduction OR > 10 degrees int. rotation and passive hip movement to neutral possible but pelvic movement occurs.

Severe: Hip held in > 10 degrees abduction AND > 10 degrees int. rotation. Unable to passively position hip in neutral.

PAIN RATING

The following guidelines should be used in order for the clinician to rate the patient's pain level.
This is an overall pain rating for the patient and should be based on the patient's presentation during the visit.

NO PAIN: Subject does not report pain and does not present with pain during or following any assessment components.

MODERATE PAIN: Subject reports knee pain in more than 3 activities AND Pain is elicited with at least two clinical tests AND Pain lasts for more than a few seconds. Clinician observes pain behaviors (facial expressions) during clinical tests, but patient allows therapist to carry out testing. Intensity of pain as well as the number of tests that are positive for pain should be considered. Pain may occur during or immediately following activities or testing.

SEVERE PAIN: Subject reports continuous pain that greatly limits all activities and pain is present prior to clinical tests and is aggravated by any clinical testing of the knee. Clinician observes pain behaviors (facial expressions, protective muscle spasm and guarding) when clinician approaches patient prior to actually carrying out the test. Therapist may not be able to carry out certain tests due to pain.

FUNCTION RATING

The following guidelines should be used in order for the clinician to rate the functional limitations (disability) of the subject.
As you will not be able to have the patient carry out all functional activities, have them perform the following activities and rate their overall function based on your observations of pain and difficulty in performing the activity.
Walking, Squatting, Jumping, Kneeling, Stepping Up, Stepping Down

Walking: Short distance (20 meters)
Squatting: full squat down and then up (can allow heels to be lifted off ground so that full knee range can be achieved)
Jumping: Standing, jump leading with both legs and landing with both legs
Kneeling: 2 point kneeling
Stepping up: leading with the affected leg then unaffected leg (8 consecutive steps if on stairway)
Stepping down: leading with the affected leg then unaffected leg (8 consecutive

steps if on stairway)

IF stairway not available use stool and have patient step up with affected limb and down with non-affected limb. Repeat leading up with the non-affected limb and step down with affected limb. Repeat this 4 times.

NORMAL: Subject carries out activities without pain and within normal limits.

MODERATELY LIMITED: Subject can complete all activities but complains of pain in at least 3 activities or clinician observes compensatory movements OR subject can not complete one activity but completes all other activities with pain or compensation in at least one.

SEVERELY LIMITED: Subject cannot complete three or more activities OR demonstrates severe compensatory movements in the majority of activities.

OTHER FINDINGS

Optional. Useful for clarification.

ASSESSMENT COMPLETED ON ONE VISIT
Yes or No

PROGNOSIS

Clinician should make a judgement of the patient's prognosis over the next month. This decision will be made based on the clinical presentation and considering the course of treatment, potential compliance of the patient and other factors that may affect the condition over the month (ie. activity level).

TREATMENT PLAN
Briefly describe treatment plan (ie. home program, regular treatment including ...)

EVALUATOR

Name and designation

FINAL EVALUATION FORM

The majority of the form is the same as the initial form, except for a few areas.
Only the areas that have changed are discussed below.

NOTE: Make ratings considering the patient's current status.

History - excluded

TREATMENT- indicate the treatment program given since initial assessment

TREATMENT COMPLIANCE

Poor - Patient did not perform any rehab activities (ie. home program) independently on a regular basis AND Patient missed treatment session frequently.

Fair - Patient attended the majority of treatment sessions and independently performed rehab program although not regularly.

Good - Patient attended all treatment sessions and independently performed rehab program on regularly.

TIME OF FINAL EVALUATION - indicate the time period since patient's initial visit for this problem, that the evaluation is being carried out.

CLINICAL CHANGE - Rate the change in patient's condition as of this visit.

PLAN - Indicate future management.